



Information Technology Standards and Guidelines Program

Data Management

Technical Standard and Guideline IT-212.02-05

February 2002

Executive Summary

This Technical Standard and Guideline (TSG) explains how to prepare a Data Management Plan for an Automated Information System. A Data Management Plan guides data management in Automated Information Systems by specifying responsibilities, approach, products, required tasks, and documentation for the systems.

This guideline contains content and format requirements for a Data Management Plan. The Chief Information Officer of the United States Patent and Trademark Office publishes this guideline. This guideline applies to all Automated Information System development projects. Compliance with this guideline is required unless the Chief Information Officer grants a waiver.

The Data Management Plan is based on the *Life Cycle Management for Automated Information Systems* manual. The Data Management Plan also supports the requirements of the United States Patent and Trademark Office policy on Data Administration, effective October 16, 1995, and the Paperwork Reduction Act.

1. PURPOSE. To guide data management for Automated Information Systems and establish standards for data management plans.
2. AUTHORITY. This publication is published under the auspices of the Chief Information Officer (CIO), United States Patent and Trademark Office (USPTO), in accordance with the USPTO Data Administration Policy, Office Administrative Directive, Automation and Automatic Data Processing 212-01, dated October 16, 1995.
3. APPLICABILITY. This guidance applies to all USPTO personnel responsible for Automated Information Systems Data Management planning, and their supporting contractors.
4. SUMMARY OF CHANGES. This TSG applies to all Automated Information Systems projects. This TSG supercedes the previous Data Management TSG, October 1995.
5. SCOPE
 - a. Compliance. Compliance with the provisions of this document is required unless a specific waiver is authorized.
 - b. Waivers. Waivers to the provision of this publication will be authorized only by the CIO, on a case-by-case basis. Waiver authorization may be delegated to the Director, Office of Data Management.
6. RECOMMENDATIONS. Recommendations concerning the contents of this technical publication should be forwarded to the USPTO OCIO Software Engineering Process Group (SEPG) and the Office of Data Management.
7. APPROVAL. The Data Management Plan shall be approved jointly by the Project Manager, the System Development Manager and the Director of the Office of Data Management.

__SIGNED_____
Douglas J. Bourgeois
Chief Information Officer

February 12, 2002_____
Date Signed

Record of Changes

CHANGE NUMBER	DATE OF CHANGE	DATE RECEIVED	DATE ENTERED	SIGNATURE OF PERSON ENTERING CHANGE

Table of Contents

1	GENERAL	1-1
1.1	INTRODUCTION.....	1-1
1.2	CONCEPTS AND DEFINITIONS.....	1-1
1.3	DATA MANAGEMENT OBJECTIVES	1-2
1.4	CONTEXT OF DATA MANAGEMENT PLAN IN THE AIS LIFE CYCLE.....	1-3
1.4.1	Prior Activities	1-3
1.4.2	Concurrent Activities	1-3
1.4.3	Follow-On Activities.....	1-3
1.5	LIST OF PRODUCTS	1-4
2	TASKS AND RESPONSIBILITIES	2-1
2.1	OFFICE OF DATA MANAGEMENT	2-1
2.1.1	Data Administration Division	2-1
2.1.2	Data Base Administration Division.....	2-2
2.2	DATA MANAGEMENT OVERVIEW	2-3
2.3	CONCEPT PHASE.....	2-5
2.3.1	Data Management Approach.....	2-5
2.3.2	Data Stewardship	2-8
2.3.3	Data Management Tools	2-16
2.3.4	Metadata Documentation Products	2-16
2.3.5	Additional Data Management Activities.....	2-17
2.4	DETAILED ANALYSIS AND DESIGN PHASE.....	2-20
2.4.1	Logical Data Model	2-20
2.4.2	Physical Data Base Design.....	2-20
2.4.3	Data Element Standardization.....	2-20
2.4.4	SGML/XML Resources	2-21
2.4.5	Data Conversion.....	2-21
2.4.6	Data Quality	2-21
2.4.7	Testing Support.....	2-22
2.4.8	Data Back-up, Logging, and Recovery Plan	2-22
2.4.9	Sensitive Data/Privacy Act Compliance	2-22
2.4.10	Mapping Matrices	2-23
2.4.11	Data Stewardship Implementation Activities.....	2-23
2.4.12	Data Base Administration Activities.....	2-24
2.4.13	Additional Tasks and Responsibilities	2-24
2.5	DEVELOPMENT PHASE.....	2-26
2.5.1	Data Stewardship Implementation Activities.....	2-26
2.5.2	Metadata Products.....	2-26
2.5.3	Data Quality	2-27
2.5.4	SGML/XML Resources	2-27
2.5.5	Testing Support.....	2-27
2.5.6	Data Base Management Activities	2-27
2.5.7	Back-up, Recovery, and Restart.....	2-27
2.5.8	Records Disposition	2-28
2.5.9	Electronic Records Management	2-28

2.6	DEPLOYMENT PHASE.....	2-29
2.6.1	Data Stewardship Implementation Activities.....	2-29
2.6.2	Data Quality.....	2-29
2.6.3	Data Base Management Activities.....	2-29
2.6.4	Information Collection Burden.....	2-30
2.6.5	Electronic Records Management.....	2-30
2.7	OPERATIONS PHASE.....	2-31
3	Content and Format.....	3-1
3.1	DOCUMENTATION STANDARDS.....	3-1
3.2	EVALUATION CRITERIA.....	3-1
APPENDIX A	BILIOGRAPHY	A-1
APPENDIX B	DATA MANAGEMENT PLAN CONTENT DESCRIPTION	B-1
APPENDIX C	DATA MANAGEMENT PLAN TEMPLATE	C-1
APPENDIX D	GLOSSARY	D-1
APPENDIX E	STANDARD DATA ELEMENT WORK SHEET	E-1
APPENDIX F	SAMPLE MAPPING MATRIX	F-1
APPENDIX G	DATA CONVERSION PLAN TABLE OF CONTENTS	G-1

1 GENERAL

1.1 Introduction

This *Data Management* Technical Standard and Guideline (TSG) provides detailed instructions for the preparation of data management plans. It applies to all United States Patent and Trademark Office (USPTO) Automated Information System (AIS) projects, as outlined in the *Life Cycle Management for Automated Information Systems* (LCM-AIS) manual.¹ Its provisions apply to all development, migration, re-engineering, maintenance, and retirement efforts for any major automated information system and limited infrastructure projects.

Guidance for AIS data modeling, data naming convention, and data element standardization do not appear in this TSG. Guidance for those activities can be found in the *Detail Design Document* and *Data Element Naming Conventions and Standardization TSG*, IT-212.03-13. Records management guidance can be found in the *United States Patent and Trademark Office Comprehensive Records Schedule*. The Data Administration Division of the Office of Data Management (ODM) is responsible for AIS project records management support for electronic, paper, and other records. Consult ODM for guidance on electronic records management and the procedures for managing Standard Generalized Markup Language and eXtensible Markup Language document resources in the *Standard Generalized Markup Language (SGML) and eXtensible Markup Language (XML) Resource Management Guidelines Technical Note*, IT-212.2-05, TN01.

1.2 Concepts and Definitions

Data management planning is an important element of system life cycle activities. It begins during the earliest phase, proceeds as requirements are defined and software is implemented, and continues until the automated system is terminated or replaced. The data management activities performed during the system development life cycle are based upon the following basic principles.

Data is recognized as a valuable resource.

Data is collected, stored, safeguarded, and used to support agency missions and business processes and decisions, making accurate and timely data an important corporate resource.

Data is defined separately from the technology used to collect and store it.

Business area data requirements are recorded clearly before designing automated data collection and storage methods, so that program needs are understood and recorded.

Accurate information about data is essential.

¹ The *Life Cycle Management for Automated Information Systems* (LCM-AIS) may be found at <http://ptoweb/ptointranet/cio/index.htm>.

Effective management of data collected by the business area requires that accurate information about data (metadata) be kept.

Common data management guidelines, methods, and tools are used.

A common approach to defining, modeling, designing, and documenting data improves data quality and makes it easier to share data among systems and offices.

A data architecture is developed and maintained.

The USPTO common data architecture is the high-level structure and organization of USPTO data. The creation of a common data architecture across the USPTO enterprise supports data sharing and information system interoperability. It defines USPTO data in a common context, addresses data quality issues, and fosters information understanding throughout the enterprise. The USPTO common data architecture is comprised of the Enterprise Data Model, data element naming conventions, standard data elements, and the Enterprise Information Repository.

1.3 Data Management Objectives

This TSG provides guidance for preparing the Data Management Plan. The objectives of the Data Management Plan support overall USPTO data policy objectives. These include:

- Recognize and promote the importance of data and information as valuable resources. This can only be achieved by proper management of the creation, use, storage, documentation, and disposition of data.
- Promote data consistency and standardization throughout the organization by developing standards for data element names, definitions, values, formats, metadata, and documentation in the Enterprise Information Repository and in data bases.
- Minimize duplication in collecting, processing, storing, and distributing data.
- Encourage and facilitate data and information sharing among USPTO business areas, the Department of Commerce, other Federal agencies, and the intellectual property community world-wide.
- Improve the quality, accuracy, and integrity of shared data resources.
- Improve data administration and access to metadata with appropriate new and existing methods, tools, and technologies.
- Reduce the cost and time to implement automated information systems.
- Implement single point of entry for data.

1.4 Context of Data Management Plan in the AIS Life Cycle

The Data Management Plan components conform to the USPTO methodology for life cycle management for an AIS. Data management planning begins in the Concept Phase and continues through the Operations Phase. The Data Management Plan is written during the Concept Phase, and is approved and placed under configuration management at the end of the Detailed Analysis and Design Phase. Change control authority to the Data Management Plan requires concurrence of System Development Management and the Office of Data Management from this point in the life cycle forward. The System Development Manager is responsible for any impact that changes to the Data Management Plan may have on the rest of the project, such as data and activity modeling, and testing. The Data Management Plan is an evolutionary document and is updated as needed during each life cycle phase.

1.4.1 Prior Activities

Data management planning begins early in the Concept Phase upon the development and approval of a Business Case and immediately following initiation of the System Boundary Agreement.

1.4.2 Concurrent Activities

In the Concept Phase, the Data Management Plan is written concurrently with other AIS documents as outlined in the Quality Assurance Plan (QAP). The Data Management Plan establishes the framework for data management development related activities in the Detailed Analysis and Design Phase. This includes data management approach, data requirements gathering, stewardship roles and responsibilities, data element identification, metadata management, model management, data management tools, records management (electronic, paper, and other records), expected information collection burden, and data privacy.

In the Detailed Analysis and Design Phase, the Data Management Plan is revised concurrently with other AIS documents to confirm the data management related activities for this phase, as well as those data management activities outlined in the Development Phase.

In the Development Phase, data management occurs concurrently with preparation of other AIS documents as outlined in the Quality Assurance Plan (QAP). It establishes the framework for subsequent Deployment Phase processes, including testing, transition, data conversion, and records management (electronic, paper, and other records).

In the Deployment Phase, data management occurs concurrently with the validation of the production environment. It establishes the framework for data base management, data integrity, and records disposition concerns during the subsequent Operations processes.

1.4.3 Follow-On Activities

Data management planning is an iterative process throughout the entire life cycle encompassing follow-on activities. In each phase, the Data Management Plan may readdress topics covered during earlier phases to refine concurrent processes.

1.5 List of Products

The resulting products required from this TSG are the Data Management Plan and, when applicable, the Data Conversion Plan. The Data Management Plan describes the data management approach that will be used and defines data related activities that will be completed for the AIS project. The Data Conversion Plan describes the details of transforming data from its original source to the newly designed target data base. A sample outline of the Data Conversion Plan is in Appendix G.

2 TASKS AND RESPONSIBILITIES

2.1 Office of Data Management

The Office of Data Management consists of the Data Administration Division (DAD) and the Data Base Administration Division (DBAD). The Data Administration Division's major tasks are performed during the Concept and Detailed Analysis and Design Phases of the project's Life Cycle Management, tapering off gradually near the end of the Development Phase. The Data Base Administration Division plays a minor role during the early phases of the project's Life Cycle Management, but actively participates and performs major tasks during the Development, Deployment, and Operations Phases of the system. The Data Base Administration Division should be consulted prior to the completion of the Detailed Analysis and Design Phase, before the data base design is completed. In addition, the Office of the Director administers the Enterprise Information Repository and the XML Resource Repository.

2.1.1 Data Administration Division

The Data Administration Division is responsible for ensuring the most cost-effective organization and use of an enterprise's data resources. The Data Administration Division is ultimately responsible for supporting AIS projects with data management support. This includes data management plan guidance, data modeling, and data element naming and standardization and facilitating the sharing of corporate data and information. The Data Administration Division is also responsible for supporting AIS projects with records management support for electronic, paper, and other records. This includes implementing sound records management and information collection practices that ensure compliance with federal laws and regulations and assisting business areas to operate more efficiently. The Data Administration Division is also responsible for supporting AIS projects with electronic records management support for storing records for effective management of the associated electronic records. The Data Administration Division responsibilities include, but are not limited to, the following activities.

- Prepare, implement, and manage data administration policies, procedures, rules, standards, and guidelines;
- Develop and maintain a consistent data architecture, including a logical Enterprise Data Model and logical data models of business areas, and support development and validation of physical models for automated systems;
- Enforce data element naming conventions on logical data models and physical data bases;
- Design, implement, and manage the data standardization process and the supporting Enterprise Information Repository of metadata;

- Prepare and approve data management plans, data models, data naming, and standard data elements;
- Manage the Data Stewardship program;
- Manage the Data Quality program;
- Advise data stewards on data quality assessments and improvements;
- Facilitate business process re-engineering through the identification of opportunities for shared data;
- Administer the USPTO records management and electronic records management programs;
- Administer the Vital Records Program;
- Administer the USPTO information collection burden program under the Paperwork Reduction Act; and
- Carry out the objectives of the Data Management policy, whether explicitly stated or not.

2.1.2 Data Base Administration Division

The Data Base Administration Division, Office of Data Management, stores, secures, and maintains data bases. This office ensures that all data bases and operational systems are running at optimum performance while providing support for new development and AIS projects. The responsibilities of the Data Base Administration Division include the following activities:

- Meet and consult with development teams, provides vendor and technical support for development teams, and create physical data bases;
- Support AIS projects by reviewing proposed changes; troubleshooting; and assisting with retrieval, web tools, special changes, strategic planning, and end user support;
- Provide support services for developing automated information systems, such as requirements analysis, data base design, and implementation and maintenance strategies for data base applications;
- Support systems programming with software maintenance, system monitoring, software management, and in-house documentation;
- Support operational maintenance with operational readiness, data base management, backup/recovery and security; and

- Monitor licensing compliance and assists the Help Desk with training and interpretation of error conditions.

2.2 Data Management Overview

This section describes the tasks and responsibilities of the USPTO employees and supporting contractors for AIS data management. Data management includes:

- Preparing the Data Management Plan;
- Defining the data management approach; and
- Defining and supporting the data management tasks.

Preparing the Data Management Plan begins with an interview with the project's System Development Manager. Initial discussion between the DAD staff and the SDM should reveal high-level information about the project methodology, tools, system architecture, data, and key players. Discussion should reveal:

- How the requirements and data are collected;
- Where the data resides and who will manage it;
- Who has programmatic control over the data;
- Who has business knowledge of the data;
- If there are any new data elements introduced or changes to existing data elements;
- If the current system's legacy data elements have not been documented or validated in the Enterprise Data Model;
- If the Extensible Markup Language (XML) techniques and tools will be used;
- If the system will be using a Commercial-Off-The-Shelf package, Object-Oriented tool or Integrated-Computer Aided Software Engineering tool;
- How the system will address records management (electronic, paper, and other records) requirements; and
- How the data will be shared with other organizations or systems.

The DAD staff will review other documents for the project such as the Concept of Operations, Business Case, System Boundary Agreement, Requirement Specifications, and the Quality Assurance Plan. The tailoring agreement that contains required and waived Life Cycle Management documents is part of the QA Plan. The Data Management Plan is part of the required Life Cycle Management documentation. The Director of the Office of Data Management holds the authority for the waiver of the Data Management Plan.

Table 2.1 provides a summary of roles and tasks for the data management planning process and each USPTO office responsible for its execution. A more comprehensive project specific listing

of roles and functions should be included in the Project Management Plan.² In Table 2.1, the same person may perform multiple functions.

Table 2.1 Summary of Data Management Activities by Roles and Functions

FUNCTION ⇒ ROLE⇓	Data Planning	Prepare and Review Data Management Plan
Program\Project Management	Consult with Data Administration on the data management approach and support the identification of data stewardship and data quality	Review and assist as needed in the development of the Data Management Plan
System Development Management	Provide support for data management development activities Coordinate data management activities for compliance with Data Management Plan Comply with Data Management Plan	Review and submit Data Management Plan for review to Quality Assurance, Program Management, Configuration Management, System Architecture and Engineering, Operations, Testing, and End-Users
Office of Data Management Data Administration	Provide support in using the Enterprise Information Repository Prepare Data Management Plan Evaluate and select data approach, methodology, and tools Provide data administration expertise for the preparation of Data Management Plan Provide XML expertise for the preparation of XML DTD/Schema development Evaluate data resources and acquisition, data quality management, records management (electronic, paper, and other), and information collection burden Provide support for the development of the logical data model Provide support for the physical data model ensuring adherence to data element naming convention and enforcement of referential integrity Provide support for standardization of data element	Conduct impact analyses, train, and generate special reports Prepare, Update, and Submit Data Management Plan to System Development Manager Evaluate and approve Data Management Plan for compliance with data administration policies and procedures Evaluate and approve XML DTD/Schema for compliance with SGML/XML Resources Management Guidelines

(Table 2.1 is continued on the next page.)

² Refer to the LCM-AIS manual for a description of the Project Management Plan.

Table 2.1 Summary of Data Management Activities by Roles (Continued)

FUNCTION ⇒ ROLE ↓	Data Planning	Prepare and Review Data Management Plan
Data Base Administration	Provide data base administration expertise in planning for development and maintenance of physical data base, including operational impact analysis, data model analysis, implementation impact analyses Plan and schedule development and operational data base implementation, and coordinate with SDM and OSAE regarding data capacity Advise on DBMS and support tools	Evaluate compliance to technical standards for physical maintenance of data resources
Quality Assurance (QA)	Evaluate Data Management Plan for compliance with LCM	Review Data Management Plan
Configuration Management (CM)	N/A	Place Data Management Plan under CM
System Architecture and Engineering	N/A	Review Data Management Plan for compliance with Technical Reference Model and Information Technology infrastructure
Operations	Provide support for data management deployment activities	Review Data Management Plan
End-User Involvement	Participate in data stewardship assignments, data resources, and acquisition Perform data quality monitoring	Review Data Management Plan

2.3 Concept Phase

The Data Management Plan is written during the early part of the Concept Phase. In the Concept Phase, the Data Management Plan will address the project data management approach, data stewardship, data management tools, metadata documentation products, and additional data management activities.

2.3.1 Data Management Approach

The data management approach has a major influence upon the success of an automation project. The data related activities, products, and decisions that must be addressed during the system life cycle constitute the data management approach. The approach includes the degree of rigor to be

followed when performing these activities and the level of formality to be used when documenting data-related life cycle products and decisions.

Implementing a data management approach is one key to the project's success. If the approach does not address data management issues, the risk of time and cost overruns for the project will increase, as will maintenance costs for the completed system and its data.

There are two criteria to determine the data management approach: the degree of data sharing and the AIS project type. Data sharing includes use of one information system's data by a second system, and utilizing the same data for multiple functions within a single information system. There are three basic ways of sharing automated data:

- Downloading/uploading information from a data store;
- Creating/replicating an external copy of a file; and
- Creating and using shared data bases.

If data sharing exists, the data management approach should include all data management activities. Following this type of approach will minimize unexpected, negative impacts upon the system and the programs it will support. Use of standardized data elements is one way of facilitating data sharing.

If an Automated Information System project is considered a stand-alone system with no data sharing, a Data Management Plan is helpful in addressing records management (electronic, paper, and other records) issues, data quality, information collection burden; ensuring the logical view of data is clear and completely addresses all business requirements; and ensuring the physical design of the data base incorporates data integrity.

2.3.1.1 Automated Information System Projects

A second criterion that influences the data management approach is the strategies used in developing an AIS. The strategy determines the required data activities and products. These include, but are not limited to, the following: Commercial-Off-The-Shelf Application, Government-Off-The-Shelf Application, Integrated-Computer Aided Software Engineering Based, Object-Oriented Based, Multiple Phased Project, Mixed Solution, and Multiple Related AISs.

- **Commercial-Off-The-Shelf Application Project**
A Commercial-Off-The-Shelf (COTS) application project requiring no supplemental code (e.g., scripts, macros, command files, configuration files, etc.) may follow a modified data management approach. No physical data base design is required. If the data base is Oracle then it should be imported into the Enterprise Information Repository. A mapping matrix is required showing all logical attributes and what they map to in the physical data base (as feasible).

Although a modified data management approach may be used for COTS, it is strongly recommended that the creation of a logical data model be considered. In this scenario, the logical data model is compared to the data contained in the potential COTS packages. A major factor in COTS package selection is the degree to which the logical data model attributes, entities, and business rules are aligned with the potential COTS. This is the major determinant in evaluating how well a COTS package meets the business requirements. Contact the Data Administration Division for instructions on tailoring the data management approach. Examples of USPTO COTS projects include the following: Enterprise Asset Management System, Patent and Trademark Assignment System, and Information Technology Facilities Management System.

- **Government-Off-The-Shelf Application Project**
A Government-Off-The-Shelf (GOTS) application project requiring no supplemental code may follow a modified data management approach. No physical data base design is required. If the data base is Oracle then it should be imported into the Enterprise Information Repository. A mapping matrix is required showing all logical attributes and what they map to in the physical data base (as feasible). Similar to COTS, it is strongly recommended that a logical data model be created. The logical model is used to evaluate whether the GOTS meets the business requirements. Contact the Data Administration Division for instructions on tailoring the data management approach. An example of an USPTO GOTS project is the Office of Human Resource System.
- **Integrated-Computer Aided Software Engineering (I-CASE) Based Project**
A project that plans to use the I-CASE based tools should include all data management activities. Products such as the logical data model and the physical data base design may be produced automatically from the I-CASE tool by the System Developer and accepted as a data management product. Products such as the logical data model should be stored in the USPTO Enterprise Information Repository. Examples of USPTO I-CASE projects include the following: Patent Application Location and Monitoring (PALM) Migration Pre-Exam Subsystem, PALM Examination and Post-Examination Subsystem, and Revenue Accounting and Management System.
- **Object-Oriented (OO) Based Project**
A project that plans to use the OO based tools should include all data management activities. Products such as the Use Case and Class diagrams may be produced automatically from this tool by the System Developer and accepted as a data management product. An example of OO Based project is Tools for Electronic Application Management.
- **Multiple Phased Project**
When a project is to be implemented in more than one phase, the Data Management approach is determined by the amount of information known about the phase to be implemented. The Data Management Plan may change as more information is revealed throughout the Life Cycle Management phases. New versions of the Data Management Plan may be necessary

as additional phases of the project are introduced. Examples of USPTO multiple phase projects are: Trademark Electronic Application Submission, Office of Human Resource System, and Trademark Image Capture and Retrieval System.

- **Mixed Solution**
A project that plans to use a combination of either I-CASE, OO, COTS, and/or custom software should include all data management activities. All metadata products (e.g., logical and physical data model, object model, data base schema) shall be stored in the USPTO Enterprise Information Repository. Contact the Data Administration Division, Office of Data Management, for guidance. Examples of USPTO mixed solutions projects are: Enterprise Address Data Component and Office of Finance Imaging System.
- **Multiple, Related AISs**
When multiple, closely related AISs are to be developed and deployed, it is possible that a single subject area Data Management Plan may be developed, addressing all data management activities that are applicable. The decision to select this option is based on complexity and manageability and will be made by the Office of Data Management. Examples of USPTO multiple related AIS projects are: Pre-Grant Publication, and Tools for Electronic Application Management.

Regardless of project type, the electronic records management requirements still apply unless waived by the Chief Information Officer. For additional details on electronic records management considerations, consult the Electronic Records Management Team Leader.

2.3.1.2 Approach Selection

Depending on the degree of data sharing and the project type, determine the data management approach to use for the project. Choose the tasks, products, and activities to be included in the data management approach. A primary objective of every automation project is to provide accurate and consistent data to the business users. Record the data management approach in the Data Management Plan early in the Concept Phase and refer to the Plan at the beginning of each phase, revising it as necessary.

2.3.2 Data Stewardship

Another major activity that the Data Management Plan addresses is the parties responsible for data. Data stewardship is defined as "...the willingness to be accountable for a set of business information for the well-being of the larger organization..."³ The data stewardship program consists of establishing teams of people (both business users and CIO staff) that have a vested interest in the content and/or structure of particular classes of data and who will assume

³ English, Larry P., *Improving Data Warehouse and Business Information Quality*, John Wiley and Sons, Inc., New York, 1999

responsibility for the preservation and quality of that data. The data stewardship program is a means to leverage the cooperation already established between CIO and the business users. It provides a means by which the data stewards can work together to ensure that data is of high quality and is useful to all who need the data to conduct their business at the USPTO. Data stewardship is a proactive approach to improving data quality and data sharing.

The goal is to simplify the flow of program data. Data should be created once and used often. Replication of data should be planned. However, currently the flow of program data is often lengthy and complex, as data is collected at various locations. Managing the complex activities, responsibilities, and relationships that arise from these data flows requires a method of determining which organizations involved in the data flows are responsible for which data-related activities. One reason that redundant and inconsistent data exists is due to the low confidence in the accuracy of data, even if that opinion is unfounded. One of the main goals of information stewardship is to improve the quality of data. As data become more accurate, reliable, timely, and complete, data sharing should increase and redundant and inconsistent data should be reduced.

2.3.2.1 Data Stewardship Roles and Responsibilities

The USPTO's data stewardship program consists of three levels of data stewardship: Business, Operational, and Technical stewards along with the Stewardship Quality Council.

- **Business Area Steward**
 - Ultimately accountable for the business data
 - Assigns the Operational Data Steward
- **Operational Data Steward**
 - Accountable for the content and business rules surrounding the business data
 - Accountable for identifying and defining electronic records from a business perspective including defining and initiating the archival processes
- **Technical Information Steward**
 - Accountable for the technical infrastructure supporting the processing of the data requirements within an AIS
 - Accountable for the management of electronic records and describing where and how to point at the subject records for the archival processes

Along with:

- **Stewardship Quality Council**
 - Resolve data related issues and data conflicts

As caretakers of specific business area data and electronic records, data stewards:

- Represent their business area in enterprise-wide data activities;

- Participate in the development, validation, and approval of data models and standard data elements involving their business area, assisting the data administrator as required;
- Participate in electronic records administration in order to deploy an electronic record keeping system where records are collected, organized, and categorized to facilitate their preservation, retrieval, use, and disposition;
- Work with the Office of Data Management to assess and improve data quality; and
- Support an enterprise-wide view of data sharing.

The data stewardship guidance presented in this section has the following objectives:

- Facilitate the assignment of responsibilities for data definition, collection, processing, storage, use, and disposition when systems and data bases are being built;
- Notify the USPTO Records Officer of system content and purpose, providing useful descriptions to facilitate official disposition instructions for the system and its records, from the National Archives;
- Ensure data meets mission and business area requirements by assigning accountability for high quality data and electronic records;
- Ensure that data definition, collection, processing, and storage methods within the organization's systems conform to applicable guidelines;
- Facilitate data sharing and reuse by clarifying roles and responsibilities involved in the definition, collection, processing, storage, and use of data; and
- Help the system developer ensure that the standard naming conventions are applied and standard data elements are used.

A discussion of each data stewardship role follows:

Business Area Data Steward

The same organization with overall responsibility for a business area's performance is responsible for ensuring that the quality of data required to support the business area is defined, collected, processed, stored, and presented in a timely and cost-effective manner. The main function of the Business Area Data Steward is to assign operational data stewards for all business functions within their organization. If a data element is used by more than one project, multiple data stewards will be assigned so that any changes to the data element can be properly coordinated.

Operational Data Steward

The Operational Data Steward is the subject matter expert from the organization or function who is responsible for the definition and collection of data and who exercises programmatic control over electronic records. Since the level to which an organization understands its data directly correlates to the level of success of any AIS, it is the role of the Operational Data Steward to define the metadata or characteristics about the data and electronic records used in their business functions, along with the derivation rules and the formats to be used for data derived from other data elements, along with security requirements. When multiple functions or organizations use the same data to support important program functions, a joint data definition effort is organized. In defining and collecting data, the Operational Data Steward(s) will be concerned with the following items:

- **MEANING** - what is the business definition of the data item (e.g., name, description, data type, length)
- **CONTENT** - what valid values, ranges, and formats the business expects
- **BUSINESS RULES** - how does the business use the data and what are its relationships with other data
- **UTILIZATION** - What are the appropriate federal and internal regulations and standards that must be met
- **SOURCE** - where in the business process did the data come from
- **TARGET** - where does the data go next
- **RETENTION** - how long must the data be available
- **DATA SENSITIVITY** - should the data be protected from unauthorized disclosure, alteration, or destruction
- **ISSUES** - are there any known problems or concerns about the data
- **THRESHOLD OF ACCEPTABILITY** - determine the minimum level of quality or integrity the data must maintain
- **DATA QUALITY**- does the data quality meet the business information needs

Technical Data Steward

The organization responsible for storage and processing of data is the Technical Data Steward. The functions carried out by the Technical Data Steward include those that have traditionally been performed by Automated Data Processing (ADP) organizations,

such as the System Development and Maintenance Managers, Data Base Administrators, and Data Maintenance Branch/Operations staff. During the development life cycle, this stewardship role will be assigned to the system development manager until the system is turned over to the operating organization. The Technical Data Stewards are responsible for the storage and processing of data. They are responsible for the following items:

- Maintain physical custody or direct control of the data, software, and other components used to store, process, communicate, and present data;
- Ensure the physical integrity of data;
- Safeguard the storage media;
- Carry out data sampling and data problem resolution;
- Act as Records Management contact; and
- Act as Paperwork Reduction Act contact.

Since both the Operational and Technical Data Stewards are responsible for ensuring that the data administration standards are met when defining and documenting data, the stewards will rely on the procedures contained in the *Data Element Naming Conventions and Standardization Technical Standard and Guideline, IT-212.03-13*, when defining data. The central information repository should also be reviewed for already existing standard definitions of data.

As an example, for the USPTO Standard Country information, the following data stewardship roles exist.

Table 2.2 Sample Data Stewardship Roles

Business Area Data Steward	Operational Data Steward	Technical Data Steward	USPTO Organization	AIS
Bob Saifer			International Liaison Staff	
	Ed Rishell		International Liaison Staff	
		Horatious Tanyi	Office of System Development and Maintenance	PALM MG Shared Objects
		Susan Shifflett	Office of System Development and Maintenance	PTAS and AHD
		Lana Chow	Office of System Development and Maintenance	ACTS
		Kathryn Tindle	Office of Data Management	

Stewardship Quality Council

The Stewardship Quality Council meets periodically to discuss data related issues of interest. Membership will consist of a Chair (Director of Office of Data Management); Vice Chair (Manager of the Data Administration Division); Coordinator (Data Administration Division Team Leader); and the Business Area, Operational, and Technical Data Stewards. The level of involvement of the council will vary depending on the volume and complexity of data issues. The Chair, Vice Chair, and Business Area Data Managers will only be involved during conflict resolution.

In addition to the data stewardship roles, the Data Management Plan will document the following roles in a system development project:

Program Sponsor

The Program Sponsor is responsible for overall project management, makes resources available to support the project, defines and validates customer requirements, and reviews progress at each LCM milestone. The Program Sponsor consults with the Office of Data Management in the selection of the data management approach. The Program Sponsor will assign the Project Manager.

Project Manager

The Project Manager is responsible for overseeing the complete effort to achieve implementation of the project. The Project Manager provides daily direction, coordination, and control for all aspects of the design, development, and deployment of the system subject to the technical direction of the Office of the

CIO and business direction of the Program Sponsor. Utilizing matrix management, the Project Manager directs the day-to-day activities of all members of the project team and ensures that all tasks and functional roles for data management planning are performed adequately in order to provide a system of sufficient quality to support program missions.

System Development Manager

Appointed by the CIO, the System Development Manager is responsible for designing, developing, and deploying an AIS under the business direction of the Project Manager. The System Development Manager ensures that the project is consistent with the agency's strategic information technology plans and is managed according to sound life cycle management principles and practices.

Maintenance Manager, Data Base Administrators, and/or Computer Operations Staff

The Maintenance Manager, Data Base Administrators, and/or the Computer Operations staff are responsible for the operation of the AIS or infrastructure system. The Maintenance Manager is responsible for the day-to-day operations of the system and for ensuring that the operational system remains consistent with the agency's strategic information technology plans.

The Data Base Administrators are responsible for ensuring that the data is available for operations as outlined in the Operational Support Plan.

The Computer Operations staff is responsible for ensuring that the infrastructure required to support the system is available in accordance with the requirements described in the Operational Support Plan.

Primary User(s)

The organization or function with the most important requirement to collect, store, and process data to perform a current or future business area function is the primary user. Sometimes the data steward organization designates the primary user, as the definer of the data the organization requires. Primary user organizations support user testing of systems during the life cycle, and ensure on-going data quality for the system data once the system is operational. If it is impossible to select a single primary user organization from among several users of the same data, then a joint data definition effort will probably result. The primary user organization is often, although not always, the data steward organization.

Ancillary User(s)

The ancillary users use data to perform business area functions, and report results to management, the Congress, and to others outside the agency. Unlike the

primary users, ancillary users must rely upon others to define and allow them access to the data.

Records Officer

The Records Officer records the dispositions (i.e., maintain, update, archive, delete or destroy) of the system software, documentation, data, output records, and backups in the USPTO Comprehensive Records Schedule. The USPTO Records Officer, Office of Data Management, serves as the Records Management contact for all projects and will oversee and coordinate the disposition scheduling process with the Operational and Technical Data Stewards.

Records Coordinator

The Records Coordinator is a person designated by the Program Office to be the point of contact for assisting with the actual disposal of records. The Records Coordinator will also assist in the retrieval of records currently archived to NARA.

2.3.2.2 Data Stewardship Implementation Activities

During the Concept Phase, contact is established between the data administrator and the project's management team to discuss the scope of the project's application and the general types of information anticipated for the application. If the project management team and the data administrator determine that multiple data steward organizations will be involved in a project, plan to involve all of these organizations beginning in the Concept Phase. Since implementation of stewardship at the data base or project level is not feasible when multiple organizations are involved, record the steward for each major function within each organization involved in the project. Ensure that Business Area, Operational, and Technical Data Stewards are recorded for all data elements. Ensure primary and ancillary users and records coordinators are recorded for the system.

When one organization defines, collects, and uses a set of data, that organization is the probable data steward. However, when data definition, collection, and use are split between multiple organizations, one faces a more difficult problem in determining stewardship. When this occurs, both of the primary user organizations should be selected as the data steward. When stewardship of data is shared between organizations, the organizations must coordinate when any permitted values or data structure is going to change so that the usability of the data is preserved for all users of this data. The Data Administration Division facilitates that coordination.

Concept Phase Data Stewardship Implementation Activities

- The Project Manager and System Development Manager will identify the organizations that will likely be the information steward organizations for the data the project requires.

- The Project Manager will assume the responsibilities of the Business Area Data Steward for the project. The Project Manager and System Development Manager will appoint individual(s) to assume the responsibilities for the Operational Data Steward. The System Development Manager will assume the responsibilities of the Technical Information Steward for the project.
- The Data Administration Division works with the appointed Operational Data Steward(s) to determine the Primary User(s), and Ancillary User(s). The Operational Steward will be involved in identifying data entity types as the project management team performs data modeling. Detailed information on data element naming is in the Data Element Naming Conventions and Standardization TSG.
- After the project management team has completed the data entity type and initial entity list, the Operational Data Steward assists in defining each data entity's meaning within the scope of the mission.

2.3.3 Data Management Tools

The automated tools that will be used during the project's life cycle to support data management activities need to be recorded in the Data Management Plan document. Explicit plans for managing the flow of metadata (information about data) between methods and tools through the life cycle should be documented. These tools include data modeling tools, data base management systems, records management (electronic, paper, and other records) tools, and information repositories. Explicitly identify any data management software to be used during the project's life cycle. Consult the USPTO's Technical Reference Model (TRM) for a current list of approved USPTO data management tools.

2.3.4 Metadata Documentation Products

This section describes the essential metadata documentation activities that must be performed during the system life cycle. While documentation is required of other life cycle products, this TSG covers only the essential metadata documentation of data requirements, physical data base designs, and production data base structures. The documentation must be completed to support later phases of the life cycle, reduce maintenance costs, and provide an audit trail from requirements to production data bases. Planning for metadata management is initiated in the Concept Phase. Careful management of the project Data Management Plan document will allow tracking of the project's data requirements through the production data base.

The System Development Managers need to plan and monitor the project's collection, use, and transfer of metadata throughout the system life cycle. Failure to do this task could delay the project, increase the cost of the project, or cause a data base to be implemented that does not meet requirements. Consult the Data Administration Division, Office of Data Management concerning the Enterprise Information Repository and XML Repository to support the project.

- **High-Level Logical Data Model**
A high-level logical data model is required for all AIS projects. Plans to develop logical models using Information Engineering Methodology (IEM) technique should be initiated during the Concept Phase. The Enterprise Data Model (EDM) must be reviewed to identify potential entities that are required for the AIS. A subset of the required entities should be extracted or copied from the EDM. This subset model shall form the foundation for data modeling in subsequent life cycle phases. The logical data model should contain at a minimum entities, entity definitions, and relationships. Entity-Relationship modeling is to be used during the Concept Phase to maintain the integrity of the Enterprise Information Architecture. The *Detail Design Document TSG* contains specific guidelines on logical modeling techniques.
- **SGML/XML Resources**
Any project developing a system that processes content and uses XML to do so, will work with the XML Registrar or XML Technical Working Group (TWG). During the Concept Phase, categories (the structure used to store resources in the XML Resource Repository) are established. For details on XML resources, please see the *Standard Generalized Markup Language and eXtensible Markup Language Resource Management Guidelines Technical Note, IT-212.2-05: TN01*.
- **Other Metadata Products**
There are other metadata documentation products that may be required for an AIS project such as the data element standardization worksheet, the XML element and logical data element mapping matrix, and the physical to logical data element mapping matrix. Reference Appendix E and F, and *Standard Generalized Markup Language and eXtensible Markup Language Resource Management Guidelines Technical Note, IT-212.2-05: TN01*. Consult the Data Administration Division for more information to tailor these needs.

2.3.5 Additional Data Management Activities

Records Management

Records Management is an integral part of AIS data management. It implements the USPTO's compliance with requirements of the Federal Records Act. The most relevant of those requirements mandate that the head of each Federal agency shall establish and maintain an active, continuing program for the efficient management of agency records. The vehicle for the management of agency records is an up-to-date records disposition schedule. Schedules are drafted with proposed dispositions by the agency Records Officer, and forwarded to and approved by the National Archives and Records Administration.

Records dispositions (i.e., maintain, update, archive, delete or destroy) are required for all AISs, including the system software, documentation, data, output records and backups.

Contact the USPTO Records Officer in the Office of Data Management, Data Administration Division to begin the disposition scheduling process.

The records management (electronic, paper, and other records) requirements for developing automated information systems are contained in the *Life Cycle Management for Automated Information Systems* manual. The Office of Data Management intranet site⁴ contains records management (electronic, paper, and other records) policy and procedures documents. Consult the Records Officer for guidance to ensure observance of these records management requirements. For additional information, see the USPTO Comprehensive Records Schedule.

Electronic Records Management

A draft checklist of requirements for electronic records management for an automated information system should be created. Consult the Electronic Records Management Team Leader for guidance and assistance. This checklist will be refined throughout the LCM to reflect more granular business area requirements.

Information Collection Burden

Under the Paperwork Reduction Act (PRA), an agency must gain approval from the Office of Management and Budget (OMB) prior to implementing any program, application, form, or system that involves collecting information, or using information collected from the public. The purpose of the PRA activity is to account for the burden hours the government places on the public and minimizing that burden as much as possible.

An agency also has an obligation to minimize the cost to the Federal Government of obtaining information necessary for the proper performance of Federal agency functions. Estimate any additional information collection burden hours imposed on Federal workers to process the information collected from the public.

The Office of Data Management, Data Administration Division, Records Officer should be consulted for guidelines on the information collection and approval process during the Concept Phase.

⁴ Office of Data Management intranet site: <http://ptoweb/ptointranet/nodm/odmfrontpage.htm>

Table 2.3 is a summary of data management activities for the Concept Phase.

CONCEPT PHASE**Table 2.3**

The Project Manager and System Development Manager works with the Office of Data Management to determine the data management approach, including metadata documentation, electronic and records management requirements, information collection burden, and roles and responsibilities of key players.	
DM Tasks	DM Products
Begin data requirements gathering Determine data management approach Identify data stewardship roles and responsibilities Identify data life cycle and tools Determine metadata documentation products Draft Data Management Plan Define high-level data architecture Establish XML Resource repository storing categories Establish preliminary records schedule Identify basic electronic records management checklist requirements Evaluate information collection burden	Draft Data Management Plan High-Level Logical Data or Object Model Draft Electronic Records Management Requirements Checklist

2.4 Detailed Analysis and Design Phase

Activities and plans regarding data management for an AIS project are expanded into more details in this phase. These activities and plans include but not limited to, logical data model; physical data base design; data element standardization worksheets; data stewardship activities; data conversion plan; data quality plan; testing support activities; data back-up, logging, and recovery plans; sensitive data and data base administration activities. Any associated metadata products should be completed before the end of the Detailed Analysis and Design Phase.

2.4.1 Logical Data Model

During the Detailed Analysis and Design Phase, the high-level logical model that was developed in the Concept Phase is refined to a fully attributed model in third normal form. The logical model contains the metadata about logical objects of the system. This includes the object's names, definitions, type, length/precision, domain information (description, range, values), format, etc. The logical model shall use standard data elements, standard abbreviations, and acronyms. Consult the *Data Element Naming Conventions and Standardization Technical Standard and Guideline, IT-212.03-13* for details on naming standards and list of standard abbreviations and acronyms. Developing the logical data model is a joint effort between the Data Administration Division and the AIS system development team. The validated logical data model and the mapping matrix should be completed before the end of the Detailed Analysis and Design Phase.

2.4.2 Physical Data Base Design

As the data model moves from logical design to physical implementation, the entity and attribute names from the logical model are transformed into physical tables and columns in the data base. The *Data Element Naming Conventions & Standardization TSG, IT-212.03-13*, contains a set of rules for naming the physical design data elements: tables, columns, foreign key columns, primary keys, indices, and referential integrity constraints. The naming convention advocated by the USPTO supports use of the common business name, which leads to "end user friendly" and consistent data element names across USPTO AISs. Refer to the *Data Element Naming Conventions & Standardization TSG, IT-212.03-13*, for more information. Consult with Data Administration for assistance and approval of all physical design names.

2.4.3 Data Element Standardization

Standardizing a data element is a required process of a project's data management. Data that are being shared by more than one system or defined as sharable in the *Detailed Design Document* are candidates for standardizing. These candidate data elements should already be defined in the project's logical model. The system developer, with the assistance of the data administrator, is responsible for providing input for the Standard Data Element Worksheet. The Standard Data Element Worksheet can be generated from Enterprise Information Repository. See Appendix E for an example worksheet using Enterprise Information Repository. Refer to the *Data Element*

Naming Conventions and Standardization Technical Standard and Guideline IT-212.03-13, or contact the Data Administration Division staff for further information.

2.4.4 SGML/XML Resources

During the Detailed Analysis and Design Phase the System Development Manager with an assistance from XML Registrar reviews the XML Resource Repository to determine if the XML resources can be used to fulfill the project needs, identifies the XML resources that can be modified to meet the project needs, and determines what new XML resources must be developed. The XML DTDs, schemas, and other XML resources should be incorporated into the project's Detailed Design Document.

2.4.5 Data Conversion

The need for a Data Conversion Plan is defined in the Life Cycle Management documents that are prepared early in the life cycle such as the Concept of Operations, AIS Project Management Plan, and the High Level Technical Architecture. If existing data will be used in the new system, for example, legacy data, Optical Character Recognition, Intelligent Character Recognition, Standard Generalized Markup Language, Hypertext Markup Language, eXtensible Markup Language or electronic imaging technology, include details of plans for data conversion or migration and the support activities.

Detailed information regarding the project's data conversion strategy shall be documented in a separate Data Conversion Plan. See Appendix G for guidance on Data Conversion Plan content. Consult the *Standard Generalized Markup Language (SGML) and eXtensible Markup Language (XML) Resource Management Guidelines Technical Note⁵*, IT-212.2-05: TN01 for details on SGML and XML resource management procedures.

2.4.6 Data Quality

The objective of defining, designing, implementing, and maintaining an AIS is to provide the information needed by an organization for its day-to-day operations and its management decision making. Data must be acceptable to its users if this objective is to be met. Generally, the acceptability of data is judged by two features: its usefulness and its quality. While these data features are well recognized, there are no standard definitions of data usefulness and data quality. This is due largely to the view that these features are subjective in nature and measurable only in a qualitative rather than a quantitative manner. However, data quality can be defined in a manner that permits its quantitative measurement.

Data quality is determined by comparing the data against a standard(s) and measuring the degree to which the data agrees with the standard(s). Data quality is defined by six attributes: accuracy, completeness, consistency, timeliness, uniqueness, and validity. By independently measuring

⁵ Consult Office of Data Management for Technical Note:
<http://ptoweb/ptointranet/nodm/odmfrontpage.htm>

the extent to which these six attributes occur and combining them for more in-depth analysis, the overall data quality can be obtained.

Much of the improvement in data quality that an organization requires will occur as a result of following the data management approach carefully. However, data quality should be measured periodically. A Data Quality Management process needs to be established. Instructions for establishing the Data Quality Management process are provided in the *Data Quality Management at the United States Patent and Trademark Office* guideline. Please contact the Office of Data Management for information about this document. The Data Quality Management process should include plans for conducting:

- Baseline Assessment - Overall assessment of the existing data quality condition
- Improvement Monitoring - Track the effect of corrective actions over time by continuously re-testing for critical defects identified in the Baseline Assessment.
- On-going Quality Monitoring - Once the data quality goals for a data environment has been achieved through improvement monitoring and corrective action, the data environment is regularly monitored to ensure that the data quality does not deteriorate.

The establishment of a data quality environment is often ignored. The following additional data quality activities are to be considered: determine data quality environment requirements; determine data quality tools; acquire data quality resources; and establish data quality environment. Data quality is not just a concern for production data bases but it extends to the backup and recovery media. Procedures are to be established for routine validation of backup and recovery media. The system's operational manager should perform routine evaluation of backup media.

2.4.7 Testing Support

Plans to support software program development and unit testing must be considered and recorded. Coordinate the plan carefully to avoid problems. Consult the *Testing Technical Standard and Guideline, IT-212.3-01* for detailed guidance. Identify data base administration support required to support testing activities.

2.4.8 Data Back-up, Logging, and Recovery Plan

Document plans for data back-up, logging, and recovery of the physical data the project creates in the development environment. Identify all data base administration support required to facilitate development activities.

2.4.9 Sensitive Data/Privacy Act Compliance

Many people think that sensitive information only requires protection from unauthorized disclosure. However, the Computer Security Act provides a much broader definition of the term "sensitive" information: "any information, the loss, misuse, or unauthorized access to or modification of which could adversely affect the national interest or the conduct of federal programs, or the privacy to which individuals are entitled under section 552a of title 5, United

States Code (the Privacy Act), but which has not been specifically authorized under criteria established by an Executive Order or an Act of Congress to be kept secret in the interest of national defense or foreign policy."

Briefly identify data security requirements and describe how they will be implemented. Responsibility for identifying sensitive data and protecting the data must be detailed in accordance with the data stewardship roles of the project. Details of data security requirements are recorded by the system development manager in accordance with the *AIS Security Planning TSG, IT-212.2-08* during the Detailed Analysis and Design Task. Document the data sensitivity level on the Data Element Standardization Worksheet.

The Privacy Act establishes certain controls over what personal information is collected by the federal government and how it is used. It applies to records if they are in a "system of records," which means they are retrieved by an individual's name, social security number or some other personal identifier. Notify the Records Officer as to whether or not the system will have Privacy Act implications.

2.4.10 Mapping Matrices

Various metadata products are developed for an AIS project. A mapping matrix is required to document the transformation between metadata products. The following matrices are to be developed when applicable: logical data model attributes to physical data elements; COTS data elements to logical data model attributes; and DTD or schema elements to either the logical data model or to object data in an object-oriented design.

2.4.11 Data Stewardship Implementation Activities

Data stewardship activities continue during the Detailed Analysis and Design Phase. The following data stewardship activities are conducted:

- The data element standardization process can begin as early as the Detailed Analysis and Design Phase. The Project Manager, System Development Manager, and Operational Data Steward work with the Data Administration Division to revise any data element worksheets that do not meet the AIS needs. The revised data element worksheet should be submitted to Data Administration for consideration. Detail information on the data element standardization process is in the *Data Element Naming Conventions and Standardization TSG, IT-212.03-13*. A sample standard data element worksheet can be found in Appendix E of this document.
- The Operational Data Steward will define attributes to describe the data entity types more fully including data type, length, and permitted values. For each attribute, a sensitivity level should be identified. For highly sensitive data elements, the Operational Steward shall identify authorized users who have access to the data element.

- The Project Manager, System Development Manager, and Operational Data Steward will work with the Data Administration Division to continue the data element standardization process by either drafting new standard data element worksheets for new data or submitting revised standard data element worksheets for potential changes to existing standard data elements. For each standard data element, record the Business Area and Operational Data Stewards. Name the organization that will support technical operation of the system and data base once it is in the Operations Phase, and record this organization as the Technical Data Steward for all associated standard data elements. Record the sensitivity level for each standard data element on the standard data element worksheet.
- The Operational Data Steward will define data quality criteria by identifying potential data quality filters.

2.4.12 Data Base Administration Activities

The AIS development team should meet with the Data Base Administration Division to discuss the following topics.

- Data Base Management System (DBMS) products and tools
- Physical data model
- Code reuse
- Data base sizing
- Estimated update volume
- Synchronization
- Shared data
- Data security
- Data Base Administration Division support required to support Development Phase activities, including physical data base creation, data base backup and recovery, data base security controls

2.4.13 Additional Tasks and Responsibilities

Additional data management tasks in the Detailed Analysis and Design Phase include information regarding information collection burden, and records management.

- Information Collection Burden
The program office should assist in preparing the information burden collection package (as required). Appropriate Federal Register Notice(s), change worksheets and/or a new information collection package should be created. Any supporting information, such as a draft of related electronic forms, should be submitted to the Records Officer.
- Records Management
The Comprehensive Records Schedule will be reviewed to determine if the project will require the development of a new records series or changes to dispositions of existing series.

- **Electronic Records Management**
Utilizing the Electronic Records Management Requirements checklist, the SDM and the Operational Records Steward should identify the entity types and data elements that will comprise the vital records for the AIS. Auditable events for use of the records should also be documented.

Table 2.4 shows data management activities for the Detailed Analysis and Design Phase.

DETAILED ANALYSIS AND DESIGN PHASE

Table 2.4

During this phase, the AIS Development Team and the Data Steward Organization work closely with the Office of Data Management when defining and refining data requirements, developing the detailed logical data model, creating the physical data model, defining the data conversion strategy, the defining data quality approach, defining electronic records management requirements, and planning data base capacity and security controls.	
DM Tasks	DM Products
Refine detailed logical data model or object model	Updated Data Management Plan
Develop data quality approach	Data Quality Plan (as required)
Develop data conversion strategy (as required)	Data Conversion Plan (as required)
Create physical data model	Update logical or object data models
Begin data element standardization	Candidate/Revised SDE worksheets
Update Data Management Plan	Mapping matrix (i.e., logical data model to physical data base, and physical data model to DTD/Schema, AIS physical data elements to physical data elements used in other systems)
Create mapping matrix	Records management retention schedule
Develop/Modify XML resources (as required)	Updated electronic records management requirements checklist
Refine records management schedule	Document Data Base Administration
Identify privacy requirements	Division support requirements
Begin information burden submission (as required)	
Refine electronic records management requirements checklist	
Identify Data Base Administration Division	
Support requirements for subsequent phase	

2.5 Development Phase

All data management products and documentation from the Detailed Analysis and Design Phase should be modified as needed in the Development Phase. These activities include completing and validating the physical data model and updating the logical data model with any additional changes, finalizing all associated mapping matrices, refining the electronic records management checklist, continuing data element standardization, analyzing data quality as required, and determining the data quality monitoring schedule.

2.5.1 Data Stewardship Implementation Activities

During the Development Phase the following data stewardship activities are conducted.

- The Operational Data Steward will identify the data elements that are not to be retained for the life of the system. The Electronic Records Keeping System will document the collection period and retention period for each of these data elements. Consult the Electronic Records Management Team Leader for further details.
- The Project Manager and System Development Manager will work with the Data Administration Division to continue the data element standardization process to ensure compliance with the technical design names for the physical data base as defined in the *Data Element Naming Conventions and Standardization TSG IT-212.03-13*.
- The Operational Data Steward will prepare data quality filters, execute baseline assessment, analyze results, and correct data errors.

2.5.2 Metadata Products

In this phase, the logical model is transformed into a physical model. The Data Administration Division will actively participate to ensure that each data element is defined accordingly to the USPTO standard physical naming conventions. In addition, the Data Base Administration Division will assist the developers with physical data definitions to best utilize the Data Base Management System. It is most unusual in this phase for physical data model/definition to change for performance or ease of coding.

If there are changes to the standard data elements that affect the metadata as defined in the worksheet, a revised worksheet for the affected data element will be prepared for approval.

Make provisions for ensuring that all required metadata is provided to the Data Administration Division, Office of Data Management for the agency's Enterprise Information Repository. The data management in this phase should include the activities detailed in the following subsections.

2.5.2.1 Data Models

The development team will complete and validate the logical and physical data models with the business user. The development team will apply the technical design naming standards to the physical data model. The Data Base Administration Division will create the data base schema in the development environment, based on the data base sizing estimate that was completed in the Detailed Analysis and Design Phase. The logical and physical data models will be loaded to the Enterprise Information Repository. Business information from the project's logical and physical data models will be migrated into the Enterprise Data Model.

2.5.2.2 Data Element Standardization

The business user will review the drafted standard element worksheet. The final draft of the standard data element worksheet should be completed before the end of the Development Phase.

2.5.3 Data Quality

The filters for data quality Baseline Assessment and Improvement Monitoring will be developed and executed. The development team and business users will analyze results to ascertain corrective action. The filters to be used to monitor data quality during the Operations Phase will be identified.

2.5.4 SGML/XML Resources

During the Development Phase, XML Resource modification and development activities continue, resulting in the final required DTDs, Schemas, and Style Sheets. XML Resources are to be tested during this phase.

2.5.5 Testing Support

The data source, data base environment requirements, and required Data Base Administration Division support will be identified for unit, integration, and acceptance testing. Consult the *Testing Technical Standard and Guideline, IT-212.3-01* for additional details. The Data Base Administration Division will closely monitor the testing procedures to identify any necessary data base changes needed for the production environment.

2.5.6 Data Base Management Activities

The activities related to the data base administration function include planning for the production data base environment, securing the production version of software, loading production data, plan and schedule implementation activities. These deployment activities should be documented in the *Production Installation Plan, Configuration Management Build Instructions, and Operational Support Plan*. Consult with the Data Base Administration Division for planning these activities.

2.5.7 Back-up, Recovery, and Restart

The procedures for back-up, recovery, and restart of production data will be defined. The procedures for ensuring that the production environment is recoverable will be documented in

the *Operational Support Plan*. Consideration needs to be given to ensure the data quality of back-up data and recovery data.

2.5.8 Records Disposition

Document plans to archive records. Document the AIS records disposition in the *Operational Support Plan*. The data disposition must comply with the *USPTO Comprehensive Records Schedule*. Consult the USPTO Records Officer for further information.

2.5.9 Electronic Records Management

The Electronic Records Management checklist will be revised accordingly. Test for quality, integrity, and security of the records. Document any transition and record conversion requirement in the *Production Installation Plan*. Consult with the Electronic Records Team Leader for additional information.

Table 2.5 shows data management activities for the Development Phase.

DEVELOPMENT PHASE

Table 2.5

The System Development team works with the Office of Data Management to complete data products, ensuring logical and physical data models adhere to the <i>Data Element Naming Conventions and Standardization</i> guidelines, ensuring XML products adhere to element and resource naming convention, and to establish the data base environment.	
DM Tasks	DM Products
Finalize Data Management Plan Complete and validate logical and physical models Apply technical design naming standard to physical model Continue data element standardization Finalize XML Resources Test XML Resources Perform Data Quality Analysis (as required) Determine Data Quality Monitoring Approach Update Electronic Records Management Requirements Checklist Identify data base environment requirements and Data Base Administration support requirements	Final Data Management Plan Final logical and physical model Revised SDE worksheets Final XML/Schema Resources Data Quality Assessment Report (as required) Updated Electronic Records Management Requirements Checklist Established data base

2.6 Deployment Phase

All the data management products from the Development Phase should be modified as needed in the Deployment Phase. The Data Base Administration Division will play an integral part in the Deployment Phase. The Data Base Administration Division will assist developers to ensure that all pieces of the system, from data base configuration and backup to data completeness, are ready for production.

2.6.1 Data Stewardship Implementation Activities

During the Deployment Phase, the Operational Data Steward will determine a schedule for monitoring data quality.

2.6.2 Data Quality

The Data Management Plan should include information regarding the data quality monitoring schedule and error correction process. This includes establishing the frequency for executing the data quality audit of the production data base and establishing error correction process as well as backup media.

2.6.3 Data Base Management Activities

The tasks and responsibilities of the Data Base Administration Division in the Deployment Phase include the following:

- **Data Loading**
Work with developer on data conversion and loading.
- **Data Base Upgrade**
Plan for future data base software/application upgrades.
- **Data Base Monitoring**
Monitor the data base and plan for growth.
- **Data Security**
Work with the Office of Information System Security to set up and manage data base user accounts, monitor user accesses to maintain license compliance, and manage file access controls.
- **Data Base Tuning**
For optimal performance, tuning should be an on-going process for both the data base configuration and application. The Data Base Administration Division works closely with

the developers and system administrator to make changes as usage, environment and data grow.

2.6.4 Information Collection Burden

Any required information collection submission should be completed. Approved OMB numbers should be applied to any associated forms.

2.6.5 Electronic Records Management

Implement other electronic records management activities per the NARA approved records schedule and as documented in the Electronic Records Management Requirements Checklist. Consult the Electronic Record Management Team Leader for additional information.

Table 2.6 summarizes key data management activities for the Deployment Phase.

DEPLOYMENT PHASE

Table 2.6

The System Development Manager works with the Office of Data Management to ensure that the data base environment is configured as specified, Enterprise Information Repository updated, and data quality meets business requirements.	
DM Tasks	DM Products
Perform Data Quality Monitoring (as required) Review XML DTD/Schema files Update Electronic Records Management Requirements Checklist Submit information collection package for OMB's approval (as required) Execute data base deployment plan	Data Quality monitoring report (as required) Revised Electronic Records Management Requirements Checklist Approved OMB form number Deployed production data base

2.7 Operations Phase

During the Operations Phase modifications are made to a production system to optimize performance and to correct deficiencies. The approach to ensure that all data management operations activities are carried out in support of maintaining the AIS is covered in this section. This section applies to changes made to the AIS after deployment. All other development activities shall adhere to the data management tasks as outlined in the preceding Concept, Detailed Analysis and Design, Development, and Deployment tasks.

The AIS maintenance team and data stewards are to work with the Data Administration staff to maintain the AIS data architecture. This includes the following tasks:

- Logical Data Model - It is the responsibility of the AIS maintenance team to identify and notify the Data Administration Division of new data requirements. The AIS maintenance team shall assist the Data Administration Division in maintaining the logical data model, including revising the AIS logical data model to reflect data architecture changes.
- Physical Data Model - It is the responsibility of the AIS maintenance team to maintain the AIS physical data model. The physical data model shall adhere to all data management policies and guidelines outlined in the *Data Element Naming Conventions and Standardization Technical Standard and Guideline, IT-212.03-13* including all applicable standard data elements and the Enterprise Data Model data structure.
- Standard Data Elements (SDE) - The AIS maintenance team is to ensure that all SDEs are used. The AIS maintenance team is to work with the Data Administration Division to revise SDEs as warranted.
- Data Quality Monitor - The Operational Data Steward(s) shall execute data quality audits and perform error corrections in accordance with the check schedule and error corrections process defined during the Development Phase. This is to be done for both the production data base and the backup media.
- Information Collection Budget - Any changes to the type of collection or the method of collecting information must be brought to the attention of the Records Officer so that appropriate change worksheets and revised submissions can be approved by the Office of Management and Budget.
- Records Management – To ensure proper disposition of agency records there must be a continuing effort to follow the guidance listed for each record series in the USPTO Comprehensive Records Schedule. The Records Officer can assist stewards in this and can provide more information on Records Management responsibilities and methodology.

- Electronic Records Management – Activities supporting Electronic Records Management in the Operations Phase include the transfer of vital records to archival storage on approved long-term storage media and any software or hardware dependant migrations of the records. These are auditable events and should be logged accordingly. The Electronic Records Management Checklist should be used to document other operational events and requirements.

The Data Base Administration Division will provide support to AISs as noted in the Operations Support Plan and related and applicable Service Level Agreements, including identifying performance issues, analyzing event logs, DBMS support, and data base recovery.

Table 2.7 summarizes key data management activities for the Operations Phase.

OPERATIONS PHASE

Table 2.7

The System Maintenance Manager works with the Data Base Administration Division to tune the data base for optimum performance, manage data base access, ensure data base physical integrity, and to troubleshoot problems. The Business Area and Operational Data Stewards work with the Data Administration Division to ensure data quality is meeting the business information needs.	
DM Tasks	DM Products
Maintain all metadata products and documentation (as required) Execute data quality monitoring filters (as required) Fine tune data base performance Follow Records Management Schedule Update Electronic Records Management Requirements Checklist (as required) Update Information Collection Package (as required)	Revised metadata products and documentation (as required) Data Quality Report (as required) Revised Electronic Records Management Requirements Checklist Information Collection Budget report

3 Content and Format

3.1 Documentation Standards

The project's Data Management Plan shall be prepared in accordance with the guidelines described in Appendix B. (See Appendix C for sample template.)

3.2 Evaluation Criteria

Below are listed evaluation criteria for the Data Management Plan to be used during informal and formal reviews.

Completeness and correctness checking:

- Data Management Approach
- Data Stewardship roles and assignments
- Data Quality Management
- Metadata Products: Data models, standard data elements, mapping matrices, and SGML/XML tags, DTD, and Schema
- Records Management
- Information Collection
- Electronic Records Management Requirements Checklist

Consistency checking:

- Data activities are harmonious across phases
- Records management activities are harmonious across phases
- Electronic Records Management activities are harmonious across phases

APPENDIX A

BIBLIOGRAPHY

American National Standard X3.138, Information Resource Dictionary System (IRDS), 1988

Barnett, Arnold, Data Administration and Data Dictionaries, 1990

Bruce, Thomas A., Designing Quality Databases with IDEF1X Information Models, New York, N.Y., Dorset House Publishing, 1992

Data Management Association, Guidelines to Implementing Data Resource Management, DAMA International, Bellevue, WA, 2001

Department Administrative Order 208-3, Major System Acquisitions for the Department of Commerce

Department Administrative Order 212-2, Management of Information Technology

Department Organization Order Series 20-14, Office of Information Resources Management

Departmental Notice 92-16, Departmental Life Cycle Management Policy, September 1, 1992

Departmental Notice 93-3, Establishing Data Administration Programs, March 22, 1993

Departmental Notice 93-4, Data Administration Policy for Department-wide Administrative System, March 22, 1993

English, Larry P., Improving Data Warehouse and Business Information Quality, John Wiley and Sons, Inc., New York, 1999

Federal Information Resources Management Regulation (FIRMR), Part 201-20.103-2 and Part 201-20.203-2

Handbook of Data Management 1998 Edition, Auerbach, 1998

Jacobson, Ivar, Object-Oriented Software Engineering: A Use Case Driven Approach, Addison-Wesley Educational Publishers, Inc., 1996

McFadden, Fred R., Modern Database Management, Addison-Wesley Educational Publishers, Inc., 1999

Martin, James, Business Area Analysis Handbook, James Martin House, England, 1989

Martin, James, Information Engineering: A Trilogy, Englewood Cliffs, N.J.; Prentice-Hall, Inc., 1989

Narayan, Rom, Data Dictionary: Implementation, Use and Maintenance, Englewood Cliffs, N.J.; Prentice-Hall, Inc., 1988

National Archives and Records Administration, A Federal Records Management Glossary, Washington, D.C., 1993

National Archives and Records Administration, Managing Electronic Records: National Archives and Records Administration Instructional Guide Series, Washington, D.C., 1990

National Institute of Standards and Technology (NIST) Special Publication 500-208, Manual for Data Administration, March 1993

Office of Management and Budget Circular A-109, Major System Acquisition

Office of Management and Budget Circular A-130, Management of Federal Information Resources, Appendix IV, Section 8.a (1) and b (2) and (6)

Paperwork Reduction Act of 1980 (Public Law 96-511)

Paperwork Reduction Reauthorization Act of 1986

Paperwork Reduction Act of 1995

Tannenbaum, Adrienne, Implementing A Corporate Repository: The Models Meet Reality, New York, John Wiley and Sons, Inc., 1994

USACE LCM Manager's Guide, Version 2.0, Appendix 8; March 31, 1996

United States Patent and Trademark Office, Office of the Chief Information Officer, Introduction to Life-Cycle Management for Automated Information Systems, Version 01.00.05, November 1994

United States Patent and Trademark Office, Office of the Chief Information Officer, Managed Evolutionary Development Guidebook, June 1993

United States Patent and Trademark Office, Office of the Chief Information Officer, User Interface Specification, Technical Guideline USPTO IT-212.4-10, 16 February, 1995

United States Patent and Trademark Office, Office of the Chief Information Officer, Detail Design Document, Technical Standard and Guideline USPTO IT-212.4-12, March, 2000

United States Patent and Trademark Office, Office of the Chief Information Officer, System Boundary Agreement, Technical Standard and Guideline USPTO IT-212.2-10, February 2001

United States Patent and Trademark Office, Office of the Chief Information Officer, Test Plan, Technical Standard and Guideline USPTO IT-212.3-x01, 19 October, 2001,

United States Patent and Trademark Office, Office of the Chief Information Officer, Data Management Plan Guideline Technical Standard And Guideline, October, 1995

United States Patent and Trademark Office, Office of the Chief Information Officer, Data Element Naming Conventions and Standardization Technical Standard And Guideline, August 2001

United States Patent and Trademark Office, Office of the Chief Information Officer, Data Element Naming Convention and Standardization Technical Standard and Guideline, November 2001

United States Patent and Trademark Office, Office of the Chief Information Officer, Records Management; Comprehensive Records Schedule; January 2001

United States Patent and Trademark Office, Office of the Chief Information Officer, Electronic Records Management Technical Standard And Guideline; November 2001, Draft

United States Patent and Trademark Office, Office of the Chief Information Officer, Total Data Quality Management (TDQM) at the United States Patent and Trademark Office, June 1997

APPENDIX B
DATA MANAGEMENT PLAN
CONTENT DESCRIPTION

DATA MANAGEMENT PLAN
CONTENT DESCRIPTION**1. GENERAL**

Briefly state the purpose of the system(s) and software to which this document applies.

1.1 Purpose of Document

This paragraph shall briefly describe the purpose of the Data Management Plan.

1.2 Scope (optional)

Briefly state the scope of the system. Consideration should be given in this section for multiple phase projects and when multiple systems are involved.

1.3 Automated Information System Project Type

Explicitly identify the type of Automated Information System (AIS) or infrastructure project. A description of AIS project types can be found in Section 2 of the *Data Management Technical Standard and Guideline, IT-212.02-05*.

1.4 Project References

This section shall list all documents referenced in this specification by number, title, revision, and date.

1.5 Terms and Abbreviations

This paragraph shall list any terms, definitions, or acronyms unique to this document and subject to interpretation by the users of the document.

2. METADATA PRODUCTS

This section describes the various USPTO metadata documentation products. The USPTO metadata documentation products include the following: logical data model, physical data model, standard data elements, SGML/XML products, metadata mapping matrices, records management, electronic records management, information collection burden, data conversion plan, and data quality plan. Section 4 identifies which metadata documentation products are applicable to the project by LCM phases.

3. IDENTIFICATION OF DATA STEWARDSHIP ORGANIZATION

Identify the key players regarding data management activities for the AIS project. At the minimum, the following roles are included: Business Area Data Steward, Operational Data Steward, and Technical Data Steward, along with User(s). Other roles can be added as appropriate and needed.

Prepare a separate *Roles and Responsibilities* appendix listing the names of the individuals performing the above functions.

4. LIFE CYCLE MANAGEMENT PHASES' ACTIVITIES AND PRODUCTS

Identify the data management tasks and the responsible parties for each data management task for each life cycle phase.

4.1 Concept Phase Activities and Products

Describe the data management approach, information collection burden, and records management activities (electronic, paper, and other records) for the project. Define the AIS data management approach, data stewardship, the data management tools, and metadata documentation products.

4.1.1 Data Management Approach

Describe the project's background and information that led to the need to develop the system. Define the project type. Identify the selected data management approach. Define the data management activities by role and function and identify the responsible parties.

4.1.2 Data Management Tools

This section shall identify the automated tools that will be used during the project to support data management activities. If different tools are used, this section shall describe the coordination efforts to keep the tools in sync.

Define plans for managing the flow of metadata through the AIS life cycle. Identify all data management software to be used.

4.1.3 Additional Metadata Products and Activities

Identify metadata products that are required for the AIS project during the Concept Phase. Determine and document plans for developing and maintaining the project's logical data model, metadata, and records management (electronic, paper, and other records).

Identify data management topics to be addressed during the Technical Review Board meeting at the conclusion of the Concept Phase.

All data products and activities must adhere to the requirements described in the *Data Management Technical Standard and Guideline, IT-212.02-05, Data Element Naming Conventions and Standardization Technical Standard and Guideline, IT-212.03-13* and *Standard Generalized Markup Language (SGML) and eXtensible Markup Language (XML) Resource Management Guidelines Technical Note, IT-212.2-05: TN01*.

4.2 Detailed Analysis and Design Phase Products and Activities

Define all data management related activities that will occur in this phase. Emphasize any changes that affect the data management approach and tools as defined in the Concept Phase. Identify the changes that relate to the project's data models and all metadata documentation products such as the Standard Data Element Worksheet, the mapping matrix, data conversion plan, data quality plan, records retention schedule, electronic records management, and Enterprise Information Repository.

Identify data management topics to be addressed during the Technical Review Board meeting at the conclusion of the Detailed Analysis and Design Phase.

4.3 Development Products and Activities

Document physical information for data base. Define the data set or files, physical records, segments, block sizes, data set allocations and physical size limits if possible. Modify the following if there are any changes since the Detailed Analysis and Design Phase.

- Testing Support (Integration and Acceptance Testing)
 - data testing strategy
 - test data acquisition
 - needed resources
- Cutover Plans
- Data Base and Metadata Management
- Data Back-up, Recovery, and Restart
- Records Retention Schedule
- Electronic Records Management
- Data Quality

Identify data management topics to be addressed during the Technical Review Board meeting at the conclusion of the Development Phase.

4.4 Deployment Products and Activities

Define plans to update and maintain the AIS metadata, data base, and Enterprise Information Repository. Determine and record the approach for ensuring all data management evaluation task activities will be carried out in support of operational assessment. This includes data quality monitoring and error correction.

Identify data management topics to be addressed during the Technical Review Board meeting at the conclusion of the Deployment Phase.

Address Deployment Phase topics that were documented during the Concept Phase and update as warranted during subsequent life cycle phases.

4.5 Operations Products and Activities

The approaches to ensure that all data management activities are performed in support of the AIS maintenance are covered in this section. This section applies to changes to AISs after deployment. The following metadata products and activities are required for the Operations Phase of the LCM: maintain metadata products, maintain electronic records audit trail, maintain data base and backup media data quality, maintain records disposition schedule, and monitor data base performance.

APPENDICES

Appendices shall be used as cited above in Section 3 and may be used to provide information published separately for convenience in document maintenance. As applicable, each appendix shall be referenced in the main body of the document where the data would normally have been provided. Appendices shall be lettered alphabetically (A, B, etc.).

APPENDIX C
DATA MANAGEMENT PLAN TEMPLATE

UNITED STATES PATENT AND TRADEMARK OFFICE



Data Management Plan

for the

[Insert Project Name]

[Insert Date]

Prepared by:
OFFICE OF DATA MANAGEMENT
DATA ADMINISTRATION DIVISION

Approved By:

Project Manager

System Development Manager

Director, Office of Data
Management

DATA MANAGEMENT
TABLE OF CONTENTS

1. GENERAL	
1.1 PURPOSE OF DOCUMENT.....	
1.2 SCOPE	
1.3 AUTOMATED INFORMATION SYSTEM CHARACTERISTICS	
1.4 PROJECT REFERENCES	
1.5 TERMS AND ABBREVIATIONS	
2. METADATA PRODUCTS	
2.1 LOGICAL DATA MODEL.....	
2.1.1 Entity Relationship Diagram.....	
2.1.2 Class Diagram	
2.2 PHYSICAL DATA MODEL	
2.3 STANDARD DATA ELEMENTS	
2.4 SGML/XML PRODUCTS	
2.5 METADATA MAPPING MATRICES	
2.6 RECORDS MANAGEMENT	
2.7 ELECTRONIC RECORDS MANAGEMENT	
2.8 INFORMATION COLLECTION BURDEN.....	
2.9 DATA CONVERSION STRATEGY	
2.10 DATA QUALITY ASSURANCE.....	
3. IDENTIFICATION OF DATA STEWARDSHIP ORGANIZATION	
3.1 BUSINESS DATA STEWARD.....	
3.1.1 Program Sponsor.....	
3.1.2 Project Manager	
3.2 OPERATIONAL DATA STEWARD.....	
3.3 TECHNICAL DATA STEWARD.....	
3.3.1 System Development Manager	
3.3.2 System Maintenance Manager, Data Base Administrators, and/or Computer Operations Staff.....	
3.3.3 User(s).....	
3.3.4 Primary User	
3.3.5 Ancillary User(s).....	
4. LIFE CYCLE MANAGEMENT PHASES' ACTIVITIES AND PRODUCTS.....	
4.1 CONCEPT PHASE ACTIVITIES AND PRODUCTS	
4.1.1 Data Management Approach.....	
4.1.2 Data Management Tools	
4.1.3 Additional Metadata Products and Activities.....	
4.2 DETAILED ANALYSIS AND DESIGN PHASE PRODUCTS AND ACTIVITIES	
4.3 DEVELOPMENT PRODUCTS AND ACTIVITIES.....	
4.4 DEPLOYMENT PRODUCTS AND ACTIVITIES.....	
4.5 OPERATIONS PRODUCTS AND ACTIVITIES	

1. GENERAL

Briefly state the purpose of the system(s) and software to which this document applies.

1.1 Purpose of Document

The purpose of the Data Management Plan for the **[insert specific project name]** project is to outline the specific data management, information collection burden, and records management (electronic, paper, and other records) activities that apply to the development and maintenance of **[insert specific project name]**. The plan identifies responsibilities for performing the tasks and when they are to be completed. The Data Management Plan (DMP) addresses specific tasks for the following life cycle phases: (1) Concept; (2) Detailed Analysis and Design; (3) Development; (4) Deployment; and (5) Operations. This document will review data management related activities to date and discuss in detail data management activities for the **[insert specific project name]** project from the Concept Phase through the Operations Phase. The DMP is an evolutionary document and is updated as needed during each life cycle phase.

1.2 Scope

Briefly state the scope of the system(s) involved in this project. The **[insert specific project name]** project requires modifications to multiple AISs. This DMP covers all AISs required to make changes for the **[insert specific project name]**. All AISs making changes to accommodate the **[insert specific project name]** shall adhere to the data management practices documented in this DMP.

1.3 Automated Information System Characteristics

The **[insert specific project name]** is:
(check all that apply)

- ☐ Commercial-Off-The-Self Application Project
- ☐ Government-Off-The-Self Application Project
- ☐ Integrated-Computer Aided Software Engineering Based Project
- ☐ Object-Oriented Based Project
- ☐ Multiple Phased Project
- ☐ Mixed Solution
- ☐ Multiple, Related Automated Information Systems
- ☐ Component Based Project

A description of the various AIS metadata products can be found in Section 2. The specific metadata products required for this Automated Information System (AIS) is outlined in Section 4.

1.4 Project References

The following documents were used as a basis for development of the [*insert specific project name*] DMP:

[List documents]

1.5 Terms and Abbreviations

This section list terms, definitions, or acronyms unique to this document and subject to interpretation by the users of the document.

[List unique terms and abbreviations]

2. METADATA PRODUCTS

Metadata is the foundation for establishing a common enterprise data architect. Data documentation is an important and indispensable part of the USPTO common data architecture. The USPTO metadata documentation products include the following: logical data model, physical data model, standard data elements, SGML/XML products, metadata mapping matrices, records management, electronic records management, information collection burden, data conversion, and data quality. This section describes the various USPTO metadata documentation products. Section 4 identifies which metadata documentation products are applicable to the [*insert specific project name*] by LCM phases. All metadata products are to be loaded into the Enterprise Information Repository.

2.1 Logical Data Model

Data modeling is the activity through which data requirements are identified and the data requirements are structured into logical data models. The objective of this task is to describe the things of interest to the business within the scope of the AIS, the relationships between them, and the business rules imposed by the program office. The logical model is developed by the Office of Data Management, Data Administration Division and is validated by the System Development Manager and business users. Depending upon the methodology used, the logical data model can be expressed in different forms. When information-engineering approach is used, the logical data model is expressed in the form of an Entity Relationship Diagram. When Object-Oriented approach is used, the logical data model is expressed in the form of a Class Diagram.

2.1.1 Entity Relationship Diagram

The entity relationship diagram (ERD) is a pictorial representation of the data needs. The ERD is a logical data model that contains entities, attributes, and relationships. All ERDs are normalized to third normal form.

2.1.2 Class Diagram

The class diagram shows the static structure of an object-oriented model: the object classes, their internal structure, and the relationships in which they participate. The class diagram should include the class name, list of attributes, list of operations, and class associations.

2.2 Physical Data Model

The physical data model is a data model adapted to meet the constraints of a specific Data Base Management System. It is a modification of the logical data model that can be tuned for performance or security purposes.

2.3 Standard Data Elements

Standardization of data elements is key to the development of an enterprise-wide data architecture. A common data architecture addresses data problems, such as inaccuracy, inconsistency, data maintenance expenses, and lack of data integration. Standardization requires

that all existing and new data be defined in a common context so that the data can be easily understood and readily shared. Each element needs to be understood in terms of what it means, then it needs to be uniquely identified, defined, named, and related appropriately. All logical and physical data elements shall conform to the USPTO data element naming conventions and standardization policies as documented in the *Data Element Naming Conventions and Standardization Technical Standard and Guideline, IT-212.03-13*.

2.4 SGML/XML Products

When Standard Generalized Markup Language/eXtensible Markup Language (SGML/XML) products are created, the following metadata products shall be generated: Document Type Definition (DTD), Style Sheet, document instances, and templates. Information contained in these documents should be linked to the AIS's logical data model, especially the DTD tags. This activity demonstrates a mapping between the data elements of the logical data model and the tags in the DTDs. XML DTDs, schemas, and other XML resources are to be included in the AIS' Detail Design Document. All versions of DTDs, Schemas, Style Sheets, samples of Document Instances, Public Entities, definitions of USPTO-established XML Namespaces, Templates, and associated documentation shall be stored in the XML Resource Repository. Consult the *Standard Generalized Markup Language (SGML) and eXtensible Markup Language (XML) Resource Management Guidelines Technical Note, IT-212.2-05: TN01* and the *Detail Design Document* Technical Standard and Guideline for specific details and examples.

2.5 Metadata Mapping Matrices

During the project life cycle, multiple metadata products are produced. A mapping matrix is required to document the transformation between metadata products. The mapping matrices include:

- between the logical data model attributes and physical data elements;
- between COTS data elements and logical data model attributes; and
- between elements in a DTD or schema to either the logical data model or to object data in an object-oriented design.

For examples see the *Standard Generalized Markup Language (SGML) and eXtensible Markup Language (XML) Resource Management Guidelines Technical Note, IT-212.2-05: TN01* and Appendix F of the *Data Management Technical Standard and Guideline, IT-212.2-05*.

2.6 Records Management

Records Management is an USPTO-wide activity. An active, effective USPTO-wide Records Management program is required by law. Such a program supports ongoing operations efficiently and facilitates the reengineering of USPTO business processes. A well-executed vital records program supports disaster recovery. The *[insert specific project name]* project must adhere to the record disposition schedule for each AIS as outlined in the *United States Patent and Trademark Comprehensive Records Schedule* and the *Records Management Handbook*.

2.7 Electronic Records Management

Attention to electronic records management is crucial to successful information technology planning and must go hand-in-hand with development of automated information systems. Electronic records shall be tracked in a record keeping system from creation. The record keeping system used must be approved in accordance with the USPTO Technical Reference Model. For specific details consult the Electronic Records Management Team Leader for guidance.

2.8 Information Collection Burden

Under the Paperwork Reduction Act, the United States Patent and Trademark Office has an obligation to minimize the cost to the Federal Government and minimizing the burden that government places on the public of obtaining information necessary for proper performance of Federal agency functions. This information collection burden extends beyond paper collection to electronic collection.

If it is anticipated that information is to be collected from a non-USPTO source, the program office is to review information collection mediums to determine if they meet the AIS needs. When new data is to be collected, it is recommended that standard data element names and data structure be considered. Using standard data elements allows uniformity for the data life cycle and supports consistency throughout the enterprise; thus reducing the ambiguity of the data element among USPTO business areas. If a standard data element does not exist, please refer to the *Data Element Naming Conventions and Standardization Technical Standard and Guideline, IT-212.03-13* for naming new data elements. The business user should work with the Office of Data Management to prepare the appropriate Information Collection Burden package for the Office of Management and Budget approval.

2.9 Data Conversion Strategy

If existing data will be used in the new system, for example, legacy data, Optical Character Recognition, Intelligent Character Recognition, Standard Generalized Markup Language, Hypertext Markup Language, eXtensible Markup Language or electronic imaging technology, include details of plans for data conversion or migration and the support activities. Consult Appendix G of the *Data Management Technical Standard and Guideline* for additional guidance.

2.10 Data Quality Assurance

The value of automated information systems is dependent upon the quality of the information they provide. However, the quality of AIS information is only as good as the data from which it is derived. With the advent of the Internet, we can no longer minimize the importance of the quality of the enterprise's data. Therefore, it is imperative that attention be given to data quality. The approach for addressing data quality should be an integral part of the AIS development, especially for data conversion and data migration projects. The approach should include data quality measuring, reporting, error correction, and on-going data quality monitoring and error correction procedures. The approach should extend beyond the operational system to the backup and recovery data. For guidance see *Total Data Quality Management (TDQM) at the United States Patent and Trademark Office* on the Office of Data Management website (<http://ptoweb/ptointranet/nodm/dmg.htm>).

3. IDENTIFICATION OF DATA STEWARDSHIP ORGANIZATION

Data stewardship encompasses the functions and responsibilities of an organization that exercises programmatic control over data on behalf of the program. The Data Stewardship organization is responsible for determining and documenting the data the system will collect and process. The data stewardship functional roles that must be performed for system development and maintenance are: Business Area Data Steward, Operational Data Steward, and Technical Data Steward, along with User(s). The following section provides descriptions of the data stewardship roles and Appendix A, Roles and Responsibilities, identifies the USPTO staff filling these roles for the [*insert specific project name*] project.

3.1 Business Data Steward

The Business Area Data Steward has overall responsibility for a business area's performance and is responsible for ensuring that the quality of data required to support the business area is defined, collected, processed, stored, and presented in a timely and cost-effective manner. The Business Data Steward role, during a system development effort, includes the Program Sponsor and the Project Manager.

3.1.1 Program Sponsor

The Program Sponsor is responsible for overall project management. The Program Sponsor makes resources available to support the AIS or infrastructure system, defines and validates customer requirements, and reviews progress at each LCM milestone. The Program Sponsor defines the data and functional requirements. The Program Sponsor appoints the Project Manager.

3.1.2 Project Manager

The Project Manager is responsible for overseeing the complete effort to implement the AIS or infrastructure system. The Project Manager provides daily direction, coordination, and control for all aspects of the design, development, and deployment of the AIS or infrastructure system under the technical direction of the Office of the Chief Information Officer (CIO) and the business direction of the Program Sponsor. Utilizing matrix management, the Project Manager directs the day-to-day activities of all members of the project team. The Project Manager must ensure that all tasks and functional roles for data management planning are performed adequately in order to provide an AIS or infrastructure system of sufficient quality to support program missions.

3.2 Operational Data Steward

The Operational Data Steward is the subject matter expert from the organization or function who is responsible for the definition and collection of data. Since the level to which an organization understands its data directly correlates to the level of success of any AIS, it is the role of the Operational Data Steward to define the metadata or characteristics about the data used in their business functions, along with the derivation rules and the formats to be used for data derived from other data elements. When multiple functions or organizations use the same data to support important program functions, a joint data definition effort is organized.

3.3 Technical Data Steward

The Technical Data Steward comes from the organization responsible for the storage and processing of data in the AIS. The functions carried out by the Technical Data Steward include those that have traditionally been performed by Automated Data Processing (ADP) organizations, such as the System Development and Maintenance Managers, Data Base Administrators, and Data Maintenance Branch/Operations staff. The Technical Data Steward has direct control of the data, software, and hardware components used to:

- Store, process, communicate, and present data;
- Ensure physical integrity of the data;
- Safeguard the storage media;
- Act as Records Management contact; and
- Acts as Paperwork Reduction Act contact.

Since both the Operational and Technical Data Stewards are responsible for ensuring that the data administration standards are met when defining and documenting data, the stewards will rely on the procedures contained in the *Data Element Naming Conventions and Standardization Technical Standard and Guideline, IT-212.03-13* when defining data. The Enterprise Information Repository should also be reviewed for already existing standard definitions of data.

During the development life cycle, this stewardship role will be assigned to the System Development Manager until the system is turned over to the operating organization at which time the stewardship role transfers to the Maintenance Manager, Data Base Administrators, and/or Data Maintenance Branch/Operations staff.

3.3.1 System Development Manager

The SDM, appointed by the CIO, is responsible for the design, development, and deployment of the AIS or infrastructure system under the direction of the Project Manager. The SDM ensures that the system is consistent with the agency's strategic information technology plans.

3.3.2 System Maintenance Manager, Data Base Administrators, and/or Computer Operations staff

The Maintenance Manager, Data Base Administrators, and/or the Computer Operations staff are responsible for the operation of the AIS or infrastructure system. The Maintenance Manager is responsible for the day-to-day operations of the system and for ensuring that the operational system remains consistent with the agency's strategic information technology plans.

The Data Base Administrators are responsible for ensuring that the data is available for operations as outlined in the Operational Support Plan.

The Computer Operations staff is responsible for ensuring that the infrastructure required to support the system is available in accordance with the requirements described in the Operational Support Plan.

3.3.3 User(s)

The organization(s) that use the system are documented in the DMP. The users of a system are categorized into two groups: Primary Users and Ancillary Users.

3.3.4 Primary User

The Primary User collects, stores, and processes the AIS or infrastructure system data. In addition, the Primary User supports user testing during the LCM Deployment phase.

3.3.5 Ancillary User(s)

These users require data used to perform business area functions, and report results to management, Congress and others outside the agency. The ancillary users must rely upon others to define and allow them access to the data.

4. LIFE CYCLE MANAGEMENT PHASES' ACTIVITIES AND PRODUCTS

This section identifies the data management tasks and the responsible parties for the data management activities for each life cycle phase.

4.1 Concept Phase Activities and Products

This section describes the data management, information collection burden, and records management (electronic, paper, and other records) for the [*insert specific project name*] project. It defines the AIS data management approach, data management tools, and identifies the required metadata documentation products.

4.1.1 Data Management Approach

The data management approach has a major influence on the success of a large automation project. The data related activities, products, and decisions that must be addressed during the system life cycle constitute the data management approach. The approach also includes the degree of rigor to be followed when performing these activities and the level of formality to be used when documenting data-related life cycle products and decisions.

A major element for determining the data management approach is the project scope. Several factors are used to determine the project scope. They include: data sharing, and project type. [*insert details regarding the specific project data management approach selection*]

The following table provides a summary of tasks by function and responsible organization for the data management planning process.

Summary of Data Management Activities by Roles

FUNCTION ⇒ ROLE⇓	Data Planning	Prepare and Review Data Management Plan
Program\Project Management	Consult with Data Administration on the data management approach and support the identification of data stewardship and data quality	Review and assist as needed in the development of the Data Management Plan.

FUNCTION ⇒ ROLE↓	Data Planning	Prepare and Review Data Management Plan
System Development Management	<p>Provide support for data management development activities</p> <p>Coordinate data management activities for compliance with Data Management Plan</p> <p>Comply with Data Management Plan</p>	<p>Review and submit Data Management Plan for review to Quality Assurance, Program Management, Configuration Management, System Architecture and Engineering, Operations, Testing, and End-Users</p>
Office of Data Management Data Administration	<p>Provide support in using the Enterprise Information Repository</p> <p>Prepare Data Management Plan</p> <p>Evaluate and select data approach, methodology, and tools</p> <p>Provides data administration expertise for the preparation of Data Management Plan</p> <p>Provides XML expertise for the preparation of XML DTD/Schema development</p> <p>Evaluate data resources and acquisition, data quality management, records management (electronic, paper, and other), and information collection burden</p> <p>Provide support for the development of the logical data model</p> <p>Provide support for the physical data model ensuring adherence to data element naming convention and enforcement of referential integrity</p> <p>Provide support for standardization of data element</p>	<p>Conduct impact analyses, train, and generate special reports</p> <p>Prepare, Update, and Submit Data Management Plan to System Development Manager</p> <p>Evaluate and approve Data Management Plan for compliance with data administration policies and procedures</p> <p>Evaluate and approve XML DTD/Schema for compliance with SGML/XML Resources Management Guidelines</p>
Data Base Administration	<p>Provide data base administration expertise in planning for development and maintenance of physical data base, including operational impact analysis, data model analysis, implementation impact analyses</p> <p>Plan and schedule development and operational data base implementation, and coordinate with SDM and OSAE regarding data capacity</p> <p>Advise on DBMS and support tools</p>	<p>Evaluate compliance to technical standards for physical maintenance of data resources</p>

FUNCTION ⇒ ROLE ↓	Data Planning	Prepare and Review Data Management Plan
System Architecture and Engineering	N/A	Review Data Management Plan for compliance with Technical Reference Model and Information Technology infrastructure
Quality Assurance (QA)	Evaluate Data Management Plan for compliance with LCM	Review Data Management Plan
Configuration Management (CM)	N/A	Place Data Management Plan under CM
Operations	Provide support for data management deployment activities	Review Data Management Plan
End-User Involvement	Participate in data stewardship assignments, data resources, and acquisition Perform data quality monitoring	Review Data Management Plan

4.1.2 Data Management Tools

The number and type of data management tools will depend upon the scope of the AIS. Data management tools are generally categorized into four groups: analysis tools, development tools, implementation tools, and delivery tools.

The [*Insert specific project name*] project shall use the following data management tools:

_____ COOL:Gen
 _____ Rochade
 _____ XML Cannon Repository Tool Suite
 _____ Quality Manager
 _____ dfPower
 _____ Rational Rose UML Tool
 _____ Versatile
 _____ DBMS (*supply name*)
 _____ Other (*supply name*)

4.1.3 Additional Metadata Products and Activities

The following metadata products and activities are required for the Concept Phase of the LCM for the [*Insert specific project name*] project:

- _____ Develop High Level Logical Data Model from Enterprise Data Model
- _____ Determine Information Collection Burden Requirement
- _____ Establish Preliminary Records Schedule
- _____ Identify Basic Electronic Records Management Requirements
- _____ Establish categories for storing resources in XML Resource Repository

During the Technical Review Board (TRB)'s review, the following data management topic(s) should be addressed (when applicable):

- Approved DMP
- Re-Use Opportunities (components and data sharing)

All data products and activities must adhere to the requirements outlined in Section 2 where applicable and the *Data Management* Technical Standard and Guideline.

4.2 Detailed Analysis and Design Phase Products and Activities

The following metadata products and activities are required for the Detailed Analysis and Design Phase of the LCM of the [*Insert specific project name*] project:

- _____ Develop Detailed Logical Data Model
- _____ Develop Detailed Physical Data Model
- _____ Identify which XML resources in the XML Resource Repository can be used
- _____ Develop/Modify XML resources
- _____ Identify Database Management System (DBMS)
- _____ Identify shared data opportunities
- _____ Create/Modify Standard Data Element Worksheet
- _____ Prepare Metadata Mapping Document(s)
 - _____ Logical data model attributes to physical data elements
 - _____ COTS data elements to logical data model attributes
 - _____ DTD or schema elements to logical data model attributes
 - _____ DTD or schema elements to object data in an object-oriented design
- _____ Prepare Information Collection Burden (ICB) OMB package
- _____ Revise/Update Records Schedule
- _____ Update or develop Electronic Records Management Requirements
- _____ Address Vital Records and Privacy Act requirements
- _____ Develop Data Conversion Strategy
- _____ Develop Data Quality Approach

During the Technical Review Board (TRB)'s review, the following data management topic(s) should be addressed (when applicable):

- Current DMP
- Validated Logical data model
- Data base scheme finalized
- XML resources established
- Metadata mapping documents completed
- ICB package status
- Records Schedule established
- Electronic Records Management Requirements identified
- System data and records retirement plans
- Re-Use Opportunities (components and data sharing)
- Number of Standard Data Element(s) (SDEs) used
- Data Conversion Plan completed
- Data Quality Approach finalized

All data products and activities must adhere to the requirements outlined in Section 2 where applicable and the *Data Management* Technical Standard and Guideline.

4.3 Development Products and Activities

The following metadata products and activities are required for the Development Phase of the LCM of the [*Insert specific project name*] project:

- _____ Finalize Database Schema
- _____ Finalize Standard Data Element Worksheet
- _____ Finalize Metadata Mapping Document(s)
- _____ Finalize Records Schedule
- _____ Implement Electronic Records Management System
- _____ Update Information Collection Burden OMB package
- _____ Produce Data Definition Language (DDL)
- _____ Test XML resources
- _____ Establish Development Database
- _____ Establish Test Database
- _____ Implement Data Conversion Strategy
- _____ Establish Data Quality Audit environment
- _____ Conduct Data Quality Audit
- _____ Perform Data Error Correction

During the Technical Review Board (TRB)'s review, the following data management topic(s) should be addressed (when applicable):

- Current DMP
- Metadata products finalized
- Metadata mapping documents completed
- XML Resource Test Report
- ICB package status
- Data Quality Audit Report
- Data Error Correction Report
- Records Schedule finalized
- Electronic Records Management System implemented
- System data and records retirement plans status
- Re-Use Opportunities (components and data sharing)
- Number of Standard Data Element(s) (SDEs) used
- Data Conversion status
- Data Quality Approach finalized

All data products and activities must adhere to the requirements outlined in Section 2 where applicable and the *Data Management* Technical Standard and Guideline.

4.4 Deployment Products and Activities

The following metadata products and activities are required for the Deployment Phase of the LCM of the [*Insert specific project name*] project:

- _____ Update Enterprise Information Repository
- _____ Update XML Resource Repository
- _____ Established Production Database
- _____ Determine Data Quality Monitoring Schedule
- _____ Determine Backup Media Data Quality Approach

During the Technical Review Board (TRB)'s review, the following data management topic(s) should be addressed (when applicable):

- Enterprise Information Repository updated
- XML resources repository updated
- Data Conversion Completed
- Production Database established

All data products and activities must adhere to the requirements outlined in Section 2 where applicable and the *Data Management* Technical Standard and Guideline.

4.5 Operations Products and Activities

The following metadata products and activities are required for the Operations Phase of the LCM of the [*Insert specific project name*] project:

- _____ Maintain Metadata products
- _____ Maintain Electronic Records Audit Trail
- _____ Perform Data Quality Monitoring and Data Error Correction of Production Database
- _____ Perform Data Quality Audit of Backup Media
- _____ Data Base performance tuning
- _____ Adhere to Records Disposition Schedule(s)

All data products and activities must adhere to the requirements outlined in Section 2 where applicable and the *Data Management* Technical Standard and Guideline.

APPENDIX A

ROLES AND RESPONSIBILITIES

ROLES AND RESPONSIBILITIES**BUSINESS DATA STEWARD**

Program Sponsor: *[insert program sponsor's name, title]*
[insert program sponsor's office name]

Project Manager: *[insert project manager's name, office name]*

OPERATIONAL DATA STEWARD

[insert specific project name] Administrator (s) *[insert operational data steward's name, office]*

TECHNICAL DATA STEWARD

System Development *[insert technical data steward's name, office]*
System Maintenance TBD
Manager

Data Base Administrator
Primary *[insert primary dba's name]*, Office of Data Management, Data Base
Administration Division

Secondary *[insert secondary dba's name]*, Office of Data Management, Data Base
Administration Division

Computer Operations Staff Office of System and Network Management, Central Computer Operation
Branch

USERS

Primary User(s): *[insert the office name of the primary user(s)]*

Ancillary User(s): *[insert the office name of the ancillary user(s) or N/A for not applicable]*

ADDITIONAL RESPONSIBILITIES

Model Manager(s):
Logical *[insert Data Administrator's name]*, Office of Data Management, Data
Administration Division

Physical *[insert Data Base Administrator's name]*, Office of Data Management, Data
Base Administration Division

Records Coordinator *[insert Records Coordinator's name, office name]*

Records Officer Susan Brown, Office of Data Management, Data Administration Division

APPENDIX D GLOSSARY

GLOSSARY

Ancillary User(s): Organization or individual that use data to perform mission functions, and reports results to management, the Congress, and to others outside the agency.

Archives (records management/all media usage): (1) The non-current records of an organization, preserved because of their continuing or enduring value. “National Archives of the United States” means those records that have been determined by the Archivist of the United States to have sufficient historical or other value to warrant their continued preservation by the Federal Government and that have been transferred to the Archivist’s legal custody and (2) the organization or agency responsible for appraising, accessioning, preserving, and making available permanent records. In the U.S. Government, the National Archives and Records Administration (NARA).

Archiving (Automated Data Processing usage): In electronic records, the process of creating a back-up copy of computer files, especially for long-term storage, backing up making a copy of a computer file for use if the original is lost, damaged, or destroyed.

Behavior: Represents how an object acts and reacts.

Collection of Information: A request for answers to identical questions, and/or for the compilation and maintenance of records.

Data: Representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by humans or automated means.

Data Administration. A high-level function that is responsible for the overall management of data resources in an organization, including maintaining enterprise-wide definitions and standards.

Data Base Management System: A software application that is used to create, maintain, and provide controlled access to user data bases.

Data Definition Language: Those commands used to define a data base, including creating, altering, and dropping tables and establishing constraints.

Data Integrity: The condition in which data is current, consistent, and accurate.

Data Steward: Organization or individual designated from within business areas to safeguard and share data. Business areas that exercise programmatic control over data have data stewardship responsibilities. As a business expert, the data steward identifies and defines data from a business perspective, participates in aspects of data administration, and facilitates data sharing outside the steward's organization. Stewards supply descriptions, definitions, and domain and security requirements for their subject data.

Data Stewardship: Functions and responsibilities of an organization that exercise programmatic control over data on behalf of the program.

Data Validation: Checking data for correctness and compliance with applicable standards, rules, and conventions.

Electronic Records: electronic, or machine-readable records, are records on electronic media. Electronic record, as defined in NARA regulations (36 CFR 1234.2), means any information that is recorded in a format that only a computer can process and that satisfies the definition of a Federal record per the Federal Records Act.

Electronic Records Management: techniques to manage automated records regardless of its format. Electronic Records Management is the broadest term that refers to electronically managing records on varied formats.

Information Collection Burden: Time, effort, or financial resources required to respond to a collection of information.

Information Repository: A knowledge base that integrates an enterprise's business information and application portfolio.

Logical Data Model: A representation of the organization's business data at its logical level including all principal subject areas, entity types, relationships and attributes of interest to the business, while maintaining technological neutrality.

Metadata: Data describing data, such as definition, source, responsible organization, format, and range of values. Metadata is organized in the form of entities, attributes, and relationships, and is generally stored in an enterprise information repository.

Object: An entity that has a well-defined role in the application domain, and has state, behavior, and identity.

Physical Data Model: A representation of the data base scheme, including detailed structure of the data and how it will be stored. Names and attributes should be enforced/carried over from the logical model.

Primary User(s): Organization or function with the most important requirement to collect, store, and process data to perform a current or future mission function.

Repository: A software tool for management of data and information that provides a mechanism for storing and processing descriptions of information and data processing resources.

State: Encompasses an object's properties (attributes and relationships) and values those properties have, and its behavior. An object's state is determined by its attribute values and links to other objects.

APPENDIX E
STANDARD DATA ELEMENT WORK SHEET

STANDARD DATA ELEMENT WORK SHEET

COUNTRY Code	
1. Submitter's Name:	Kathryn Tindle
2. Submission Date:	July 20, 2000
3. Phone Number:	703-308-7395
4. Office:	Office of Data Management, Data Administration Division
5. Automated Information System(s):	Enterprise Address Data Component Application Capture and Review System PALM MG Pre-Exam Patent Application Capture and Entry PCT Operations Workflow and Electronic Review Revenue Accounting Management System Electronic Application Compliant System
6. Common Business Name:	Country Code
7. Candidate Data Element Name:	COUNTRY Code
8. Data Element Description:	The code that represents the officially designated abbreviation for a country according to the International Organization for Standardization (ISO) under International Standard 3166-1.
9. Data Element Disposition:	Standard
10. Type:	Alphabetic
11. Length/Precision:	CHAR 2
12. Format:	N/A
13. Alias(es):	N/A
14. Domain Description:	N/A
15. Domain Range:	N/A
16. Domain Values:	Please refer to the International Standard 3166-1 from the International Organization for Standardization (ISO). These approved codes for use at USPTO are stored in the PTO STND COUNTRY group based on the PTO STND ISO 3166-1 table. Note: The World Intellectual Property Organization works closely with ISO so the WIPO country codes are the same as the ISO codes for countries.
17. SGML/XML Tag:	CTRY, B130
18. Standard Abbreviated Programming Name:	N/A

19.	Existing Programming Name(s):	CTRY_CD COUNTRY_CODE
20.	Authority:	ISO 3166-1 WIPO ST.3
21.	Sensitivity Level:	Low
22.	Data Structure Reference:	N/A
23.	Unit of Measure:	N/A
24.	Integrity Rules:	Since WIPO ST.3 includes both country codes and international patent organization abbreviations, only the country codes portion of the list are allowed into this address domain.
25.	Model Reference(s):	1) EADC_R01_V02_DANL_A_X of COOLCSE1 2) ICT2_R02_V03_CMPI_A_O of COOLCSE1 3) ICT2_R02_V03_CMPS_A_O of COOLCSE1 4) IEA1_R01_V06_CMPS_A_O of COOLCSE1 5) IEA1_R01_V07_CMPI_A_O of COOLCSE1
26.	Other Source:	APS Green Book Page 3 EFID Dictionary Field 3.32c, Page 151 Field 3.33e, Page 158 Issued Patents Data Dictionary Page A-34 PALM Data Dictionary Page 420 USPAT Reload Data Base Specification Page D-41
27.	Mission Area Reference:	Dissemination, Patents, Trademarks
28.	Business Area Reference:	Dissemination, Patents, Trademarks
29.	Business Data Steward:	Name: Robert Saifer Title: Director, International Liaison Staff Phone: Crystal Park Three, Suite 902 Address: 703-308-6853
30.	Operational Data Steward:	Name: Ed Rishell Title: International Liaison Staff Phone: 308-6867 Address: Crystal Park Three, Suite 902
31.	Technical Data Steward:	Name: Phong Ly Title: Manager, System Development Infrastructure Phone: 305-8719 Address: Crystal Park Three, Suite 402

APPENDIX F
SAMPLE MAPPING MATRIX

SAMPLE MAPPING MATRIX

Existing TRAM System						TRAM Replacement System			
Data Set Name	Volumes	Item Name	Business Name	Length	Type	Disposition	Entity Type	Attribute	Notes
Application Mark	2,197,152	AM-ADDR.PTODB1	Correspondence Address (occurs 5)	40	ALPHA	Modeled	Address	All Address attributes; when Functional Role Type = Correspondent	
		AM-ATTY-DKT-NUM.PTODB1	Attorney Docket Number	12	ALPHA	Modeled	Application SubType	Attorney Docket Number	
		AM-BTCH-NUM.PTODB1	Batch Number	3	ALPHA	Not Modeled: The decision on whether to process case files in batches and what data is needed to support it will be made during design.			
		AM-CHRG-TO-LOC.PTODB1	Charge to Location	3	ALPHA	Modeled: Through relationship between Case File and Case File Location			TBD: through "charge to" relationship to location?
		AM-CHRG-TO.PTODB1	Charge to Employee	5	NUMBER	Modeled: Through relationship between Case File and Employee			TBD: through "charge to" relationship to functional association?
		AM-CLS-CT-ACTV.PTODB1	Class Count - Active	2	NUMBER	Not Modeled: The count of active classes can be derived through Goods and Services entity type or can be added as a design attribute if performance warrants			

AM-CNCL-CD.PTODB1	Cancellation Code	1	ALPHA	Modeled	Registration SubType	Cancellation Code	Stored as 1 character vs. 5 character because first 4 characters are always "CNCD"
-------------------	-------------------	---	-------	---------	----------------------	-------------------	--

APPENDIX G
DATA CONVERSION PLAN
TABLE OF CONTENTS

DATA CONVERSION PLAN

TABLE OF CONTENTS

1 INTRODUCTION

- 1.1 Purpose
- 1.2 Scope
- 1.3 Project References

2 ANALYSIS

- 2.1 Data File Analysis
- 2.2 Assessment of Conversion Complexity
- 2.3 Identify Rules and Edit Criteria for Source Data
- 2.4 Identify Conversion Requirements

3 SYSTEM ENVIRONMENTS

- 3.1 Source System(s) Environment
 - 3.1.1 Hardware Configuration
 - 3.1.2 Software
 - 3.1.3 File/Data Base Structure
 - 3.1.4 Relevant Source Data Files
- 3.2 Staging System Environment
 - 3.2.1 Hardware Configuration
 - 3.2.2 Software
 - 3.2.3 File/Data Base Structure
 - 3.2.4 Relevant Source Data Files
- 3.3 Target System Environment
 - 3.3.1 Hardware Configuration
 - 3.3.2 Software
 - 3.3.3 File/Data Base Structure
 - 3.3.4 Relevant Source Data Files

4 DATA CONVERSION

- 4.1 Conversion Methodologies
 - 4.1.1 Extraction and Transformation of Source Data
 - 4.1.2 Preparation, Conversion and Loading of Staging and/or Target Data
 - 4.1.3 Synchronization
- 4.2 Conversion Management
- 4.3 Conversion Activities
 - 4.3.1 Standardize Field Definitions
 - 4.3.2 Source to Target Data Element Mapping Matrix
 - 4.3.3 Load Data Process
 - 4.3.4 Coordinate Processing Between the Current and Target Environments
 - 4.3.5 Site Preparations
 - 4.3.6 Conversion Process Testing
 - 4.3.7 Execution of Conversion
 - 4.3.8 Data Quality
 - 4.3.8.1 Data Quality Approach

- 4.3.8.2 Data Quality Environment
 - 4.3.8.3 Data Quality Assessment
 - 4.3.8.4 Error Correction
- 4.3.9 Records Management
 - 4.3.9.1 Records Disposition
 - 4.3.9.2 Records Schedule
- 4.3.10 Backup and Recovery
- 4.4 Conversion Schedule
- 4.5 Resources Required
- 4.6 Constraints

5 RISKS

- 5.1 Risks
- 5.2 Mitigation Strategies

APPENDIX A	SOURCE TO TARGET DATA ELEMENT MAPPING MATRIX
APPENDIX B	DATA QUALITY BASELINE ASSESSMENT REPORTS